REMARKS

Claims 1-7 and 9-14 are pending in the instant application and Claims 8, and 15-21 have been amended. No new claims have been and no new matter is introduced by the amendments to the claims. Applicants submit, without acquiescing in the Examiner's positions and solely in the interest of expediting prosecution of the instant application, that the amendments presented above adopt the examiner's suggestions for resolving the claim objections and for placing the claims in allowable form and thus may be entered after final rejection under 37 CFR §1.116(a). Support for the amendments to the claims can be found throughout the specification and drawings, particularly Example 2, paragraphs 188-201, and Example 3, paragraphs 202-273 of the application.

Claim Objections

Claim 10 is objected to by the Examiner due to its recitation of "anthocyan" in line 4 and repetition of "line" in line 2. Applicants respectfully submit that Claim 10, as currently amended, conforms with the Examiner's comments and thus withdrawal of this objection is respectfully requested.

Claims Rejections

Rejection Under 35 U.S.C. § 112, second paragraph

Claims 10 and 11 stand rejected under 35 U.S.C. § 112, second paragraph, as indefinite in their recitations of "R-nj" and "R-nj::Ac." In particular, the Examiner argues that the Applicant has not cited sufficient examples in support of their argument that "R-nj" and "R-nj::Ac" are well known in the art and therefore asks that the be recited in their full-form.

Applicants reiterate that R-nj is term that is well known in the art. In addition to the original disclosure of the R-nj locus in the reference included with the Applicants' previous response, Applicants direct the Examiner's attention to the following references that discuss the R-nj chromosomal locus: Dellaporta et al., Plenum Press:263-282 (1988), "R-navajo (R-nj): Isolated from the Cudu strain, this stable allele is distinguished most obviously by a solid patch of aleurone pigment in the crown region of the kernal."; Ludwig et al., PNAS, 86, 7092-7096 (1989), "Recently, the R-nj gene was cloned by transposon tagging and shown to be homologous with other members of the R gene family. In the study reported here, the R-nj clone was used to isolate genomic and cDNA clones of the Lc member of the R gene family." (internal citation ommitted); Li et al., Genetics 159:1727-1740 (2001), "R-nj:Cudu (Navajo) controls pigmentation in the crown of the kernal, the root, coleoptile, scutellar node, silk and anthers..."; and Walker et al., Genetics, 146(2):681-93 (1997), "At the r locus, both simple and complex gene arrangements are found. In simple r alleles, e.g., R-nj, a single gene is present..." (internal citation omitted). Both Dellaporta et al. and Ludwig et al. were provided in the Information Disclosure Statement submitted October 14, 2004, and courtesy copies of those references along with copies of Li et al. and Walker et al. are attached for the Examiner's convenience. In light of the extensive use of the term "R-nj" in the art without any additional explanation of its meaning, Applicants respectfully request the Examiner reconsider this rejection.

Applicants also respectfully direct the Examiner's attention to http://www.maizegdb.org/maize_nomenclature.php ("A Standard For Maize Genetics Nomenclature" From MNL 69:182-184 (1995), as updated Sep 1996; Apr 2000; Apr 2002; Oct 2006), a hard copy of which is attached for the Examiner's convenience, which describes in detail the nomenclature generally employed in this field and which the Applicants have

incorporated into the instant claims. Specifically, Applicants note that section 7 entitled "MUTATIONS RESULTING FROM TRANSPOSABLE ELEMENT INSERTIONS" states that "When the transposable element insertion [Ac, Ds, Spm(En), dSpm(I), Mu1..MuX, etc.] is known, it is suggested that this be indicated by a double colon following the allele as wx-m1::Ds1." Accordingly, Applicants recitation of "R-nj::Ac" employs the standard nomenclature used within the field for expressing an allele where an Ac element is inserted into the R-nj locus. In light of the amendments to the claims and the well known nature of the recitations in the art, Applicants respectfully request withdrawal of the instant rejection.

Claims 10 and 11 stand further rejected under 35 U.S.C. § 112, second paragraph, as indefinite in their recitation of "including the embryo (R-nj::AC allele)", which the Examiner argues is confusing.

Without acquiescing in the Examiner's position that the recited phrase is indefinite, and solely in the interest of expediting prosecution of the instant application, Applicants have amended Claim 10. Applicants respectfully submit that the amended recitation of "R-nj::Ac allele" merely represents an allele including an active Ac sequence within the R-nj chromosomal locus as discussed above.

Claims 10 and 11 stand further rejected under 35 U.S.C. § 112, second paragraph, as indefinite in their recitation of "Ac located in R-nj chromosomal region." Specifically, the Examiner argues that the "R-nj chromosomal region" is not defined and that the metes and bounds of "located" are not defined.

Without acquiescing in the Examiner's position that the recited phrase is indefinite, and solely in the interest of expediting prosecution of the instant application, Applicants have amended Claim 10. Applicants respectfully submit that the bounds of the "R-nj

chromosomal locus" are well known in the art as evidenced by Ludwig et al., PNAS, 86, 7092-7096 (1989). Furthermore, applicants have amended the claim such that the Ac sequence is within that locus, which clearly defines that the Ac sequence is found within the bounds identified above.

Rejection Under 35 U.S.C. § 112, first paragraph (Enablement)

Claims 1-7 and 9-14 stand rejected under 35 U.S.C. §112, first paragraph, as containing subject matter which was not described in the specification in such a way as to enable one of skill in the art which it pertains, or with which it is most nearly connected, to practice the invention commensurate in scope with these claims. Specifically, the Examiner argues that the instant claims are directed to mobilizable sequences from any type of transposon, while the application only enables mobilizing sequences derived from maize transposons.

Without acquiescing in the Examiner's position, and only in the interest of expediting prosecution of this application, Applicants have amended the claims such that they are directed to mobilizable sequences derived from maize transposons. In light of the foregoing, Applicants contend that the claims, as currently amended, are fully enabled by the specification and thus withdrawal of the instant claims is respectfully requested.

Rejections Under 35 U.S.C. § 102(b)

Claims 1-4 and 6-8 stand rejected under 35 U.S.C. § 102(b) as anticipated by Perez et al., WO 98/38323. In particular, the Examiner alleges that the method for obtaining a transgenic monocotyledon plant containing a gene of interest that is free of ancillary sequence, contains each and every limitation of the rejected claims.

Applicants respectfully traverse the foregoing rejection and assert that the instant claims are not anticipated by Perez et al. A proper rejection of the claims requires the Examiner to show that each and every element as set forth in the claim is found, either expressly or inherently, in the asserted reference.[2] This has not been done.

Although Perez et al. does disclose a method for the introduction of a gene of interest and the removal of ancillary sequences, Perez et al. does not teach that the active transposase should or could be located within a phenotypic marker for excision as in the rejected claims. In fact, the transposase coding sequence described in Perez et al. is located near a GUS marker, but not within that marker. Furthermore, it is indicated that Ac transposase used in Perez et al. does not contain its 5' end, which includes the Ds element essential element for auto-excision activity.

The Examiner concedes the above-described differences between the cited art and the instant invention, but argues that even in light of these differences, Perez et al. "discloses all of the active method steps that are recited in the instant claims 1-4, and 6 to practice the instantly claimed method..." (See Page 7 of the instant Office Action). In particular, the Examiner asks the Applicants attention be drawn to "Example 5 (page 7, right column) and Figure 3, wherein it is clearly disclosed that the vector which provided the source of transposase is created by deleting GUS coding region." In response Applicants respectfully point out that the transposase identified in Example 5 and Figure 3 is not operable in the context of the active method steps recited in the instant claims. Specifically, Applicants note that step (d) of Claim 1 includes a selection step that requires the expression of a phenotypic marker for excision that can only occur if the endogenous transposase is capable of auto-excision:

[2] See, Verdegaal Bros. v. Union Oil Co. of California, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987); "A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.".

(d) selecting a maize plant or cell from the F1 generation having the endogenous active transposase excised from the sequence encoding a phenotypic marker for excision based on expression of the phenotypic marker for excision ~~expression~~ to identify maize plants containing the coding sequence of interest (i) but lacking the ancillary selection marker coding sequence (ii); and

As pointed out above, the transposase described in Perez et al. is incapable of such auto-excision, and is clearly identified as lacking that ability in Figure 3 by the use of the term "Fixed Transposase." Thus the Examiner has not identified any step in Perez that correlates to this selection of the F1 generation based on the expression of a phenotypic marker for excision. While it appears the Examiner is attempting to argue that the GUS gene may provide this function, Perez et al. do not teach or suggest using GUS expression as a method of determining whether excision has occurred. Furthermore, Perez et al. could not as the transposase is itself incapable of auto-excision and the GUS gene is rendered non-functional by the insertion of the transposase.

Claims 1-3 and 6-8 stand rejected under 35 U.S.C. § 102(b) as anticipated by Yoder et al., WO 98/38323. In particular, the Examiner alleges that the method for obtaining a transgenic monocotyledon plant containing a gene of interest that is free of ancillary sequence, contains each and every limitation of the rejected claims.

Applicants respectfully traverse the foregoing rejection and assert that the instant claims are not anticipated by Yoder et al. As pointed out above, a proper rejection of the claims requires the Examiner to show that each and every element as set forth in the claim is found, either expressly or inherently, in the asserted reference. This has not been done.

As was the case of Perez et al., Yoder et al. fail to teach the inclusion of an active transposase within a phenotypic marker for excision as required by the rejected claims. Since, as above, each and every element of the rejected claims is not found in Yoder et al, as required under 35 U.S.C. § 102(b), reconsideration and withdrawal of the instant rejection is requested.

Rejections Under 35 U.S.C. § 103(a)

Claims 9 and 13-14 stand rejected under 35 U.S.C. § 103(a), as unpatentable over Perez et al., in light of Ishida et al., Nature Biotech. 14(6):745-750, 1996). In particular, the Examiner argues that it would have been prima facie obvious to one skilled in the art at the time the invention was made to modify Perez et al.'s method for obtaining a transgenic corn plant containing a gene of interest that is free from foreign ancillary sequence (as discussed above) by using inbred line A188 taught by Ishida et al. Applicants respectfully point out that the examiner has not met his burden of establishing a prima facie case of obviousness. The Examiner must establish that the cited art teach or suggest each and every limitation of the pending claims. [3] This has not been done.

As outlined above, Perez et al. fails to teach or even suggest the use inclusion of the transposase within a phenotypic marker for excision. Furthermore, Perez et al. fails to teach or even suggest a transposase having auto-excision activity due to the lack of the appropriate 5' sequence necessary for auto-excision. Applicants submit that Ishidia et al. fails to cure the deficiency of Perez et al., as it also fails to describe a system wherein a transposase is included within a phenotypic marker for excision. In light of the foregoing, Applicants respectfully

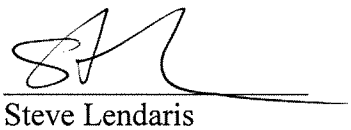
[3] In re Royka, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). "All words in a claim must be considered in judging the patentability of that claim against the prior art."; See also M.P.E.P § 2143.03

submit that the Examiner has failed to establish a prima facie case of obviousness and therefore withdrawal of the instant rejection is respectfully requested.

Conclusion

Entry of the foregoing amendments and remarks into the file of the above-identified application is respectfully requested. An early allowance is earnestly sought. To expedite allowance of this application, the Examiner is invited to telephone the undersigned if the Examiner believes a telephone call would be helpful in advancing prosecution.

Respectfully submitted,



Steve Lendaris
PTO Reg. No. 53,202
Attorneys for Applicants
BAKER BOTTS L.L.P.
30 Rockefeller Plaza
New York, NY 10112--4498
(212) 408-2500

Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 7092–7096, September 1989
Genetics

Lc, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the *myc*-homology region

(chalcone synthase/DNA-binding protein/regulatory gene/*Zea mays*)

STEVEN R. LUDWIG*, LEDARE F. HABERA*, STEPHEN L. DELLAPORTA†, AND SUSAN R. WESSLER*‡

*Department of Botany, University of Georgia, Athens, GA 30602; and †Department of Biology, Yale University, New Haven, CT 06511

Communicated by Barbara McClintock, June 23, 1989 (received for review May 5, 1989)

ABSTRACT **Previous studies have suggested that the *R* locus of maize is responsible for determining the temporal and spatial pattern of anthocyanin pigmentation in the plant. In this report we demonstrate that three members of the *R* gene family, *P*, *S*, and *Lc*, encode homologous transcripts 2.5 kilobases in length. The structure of one *R* gene, *Lc*, was determined by sequencing cDNA and genomic clones. The putative Lc protein, deduced from the cDNA sequence, is composed of 610 amino acids and has homology to the helix–loop–helix DNA-binding/dimerization motif found in the L-*myc* gene product and other regulatory proteins. It also contains a large acidic domain that may be involved in transcriptional activation. Consistent with its proposed role as a transcriptional activator is our finding that a functional *R* gene is required for the accumulation of transcripts of at least two genes in the anthocyanin biosynthetic pathway. We discuss the possibility that the diverse patterns of anthocyanin pigmentation conditioned by different *R* genes reflect differences in the *R* gene promoters rather than their gene products.**

The anthocyanin biosynthetic pathway of maize has proven to be an ideal system for understanding genetic interactions between regulatory and structural genes (for review see ref. 1). The presence or absence of pigment is a sensitive nonlethal phenotype that has been exploited in the identification of at least 10 loci required for expression of the wild-type purple color (1). Transposable elements at some of these loci have facilitated the cloning of regulatory and structural genes by transposon tagging (2–6). Molecular characterization of one presumed regulatory locus, *c1*, has revealed that the *C1* gene product is required for transcription of structural genes in the anthocyanin biosynthetic pathway (6). In addition, the putative C1 protein has the structural features of a transcriptional activator and shares sequence homology with the proteins encoded by the *myb* protooncogenes (7).

Whereas the *C1* gene product is required for aleurone pigmentation, the product of another presumed regulatory gene, *R*, is required for the pigmentation of all plant tissues (8). Although some maize strains have only a single *R* gene, others have as many as four (9). It is for this reason that we refer to these genes as the *R* gene family. This family includes nearly 100 alleles that have been extracted from maize strains found in diverse geographic locations (10–12). When crossed into a common genetic background, expression of these genes can be distinguished by differences in the spatial distribution of pigmentation in the plant (9–11).

The pattern of pigmentation displayed by a particular plant reflects the combined expression of all *R* family members that it contains (13). For example, the "standard" *R* locus is responsible for pigmentation of the aleurone, anthers, and coleoptile. This phenotype is due to the expression of two tightly linked members of the *R* gene family, *S* and *P* (14). While *S* pigments the aleurone of the kernel, *P* pigments the anthers and coleoptile of the plant. Gene duplication and divergence of the "standard" *R* locus in some strains are thought to have produced another member of the *R* gene family, *Lc* (15). The *Lc* gene, which is located 2 map units distal to the *R* locus, conditions the pigmentation of midrib, ligule, auricle, glume, lemma, palea, and pericarp tissues (15).

Although the function of the *R* gene family is unknown, prior genetic and biochemical studies indicated that the *R* gene is required for the enzymatic activities encoded by the *a1*, *c2*, and *bz1* loci (16–18). Recently, the *R-nj* gene was cloned by transposon tagging and shown to be homologous with other members of the *R* gene family (19). In the study reported here, the *R-nj* clone was used to isolate genomic and cDNA clones of the *Lc* member of the *R* gene family. Evidence is presented that the putative protein product of *Lc* has features of a transcriptional activator and is required for the accumulation of transcripts of two structural genes in the anthocyanin biosynthetic pathway.§

MATERIALS AND METHODS

Maize Stocks. All *Zea mays* seeds used in this study contained the common genetic background of the inbred strain W22 and were kindly provided by Jerry Kermicle (Laboratory of Genetics, University of Wisconsin, Madison). The anthocyanin genes present in W22 are *C2*, *A1*, *A2*, *C1*, *pl*, and *B-b*. *B-b* was replaced by *b* for the RNA analysis of *R*, *C2*, and *A1*.

Genomic DNA and RNA Isolations and Hybridization Analyses. Maize genomic DNA was isolated from young leaf tissue, digested with restriction enzyme, transferred to nitrocellulose, and hybridized as described (20). Total RNA was isolated according to Fedoroff *et al.* (21) from kernels 35 days after pollination (dap), from male spikelets of tassels that had shed pollen for 1 day, and from female spikelets 10 dap with a 2-day exposure to light at 8 dap. Poly(A)$^+$ RNA was purified by oligo(dT)-cellulose chromatography (22), fractionated in 1.2% agarose/formaldehyde gels, and transferred to Magnagraph membranes (Fisher). The filters were prehybridized for 4 hr at 65°C in 5% SDS/0.33 M sodium phosphate, pH 7.0/0.1 M EDTA containing heparin at 0.15 mg/ml. Hybridizations were performed for 18 hr at 65°C in the prehybrization solution. Filters were washed twice for 45 min at 68°C in 15 mM NaCl/1.5 mM sodium citrate, pH

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: dap, days after pollination.

‡To whom reprint requests should be addressed.

§The sequence reported in this paper has been deposited in the GenBank data base (accession no. M26227).

7.0/0.5% SDS. DNA probes were labeled with [α-^{32}P]dATP by either primer-extension (23) or nick-translation (24).

Genomic and cDNA Cloning. Genomic and cDNA clones of *Lc* were isolated as follows. For the 3.7-kilobase (kb) *Hin*dIII fragment, total genomic DNA was digested to completion with *Hin*dIII, ligated to *Hin*dIII-digested Charon 35 arms, packaged *in vitro*, and plated on K803 cells (25). Of 5×10^5 phage screened with the pR-nj:1 insert (19), 1 recombinant phage contained the 3.7-kb *Hin*dIII genomic fragment. A cDNA library in λgt10 was constructed using a Pharmacia cDNA kit and 5 μg of poly(A)$^+$ RNA from female spikelets isolated 10 dap. The library was screened with the 3.7-kb *Hin*dIII genomic *Lc* clone by standard methods (26). Of 3×10^5 recombinant phage screened with the 3.7-kb *Hin*dIII fragment, only 1 phage containing a 2.5-kb insert was homologous to the genomic clone. To isolate the 6-kb *Hin*dIII genomic fragment, *Hin*dIII genomic fragments of 5–7 kb were ligated into the *Spe* I site of λZAP (Stratagene), packaged, and plated. Of 2×10^5 recombinant phage screened with the 3′ end of the cDNA (positions 880–1772), 12 phage contained the 6-kb *Hin*dIII fragment.

Genomic and cDNA inserts were subcloned into plasmid pUC119 (27). Overlapping subclones and unidirectional deletion clones were isolated (28) and sequenced by the dideoxy method of Sanger *et al.* (29).

Primer-Extension and RNase Protection Analysis. The start of transcription was determined by primer-extension according to the method of Dunsmuir *et al.* (30). Five micrograms of poly(A)$^+$ RNA and 0.1 pmol (7.5×10^5 cpm) of the ^{32}P-labeled oligodeoxynucleotide 5′-CGTGAACCGGCG-GACGAGGG-3′ were hybridized at 55°C for 3 hr. The primer was extended for 45 min at 37°C with avian myeloblastosis reverse transcriptase. RNase protection experiments were performed according to Promega. A 1.6-kb *Hin*dIII–*Sph* I *Lc* genomic fragment was subcloned into pGEM-4Z (Promega). The plasmid (p259) was linearized with *Hin*dIII and transcribed *in vitro* with SP6 polymerase. The labeled RNA (1.5×10^5 cpm) was added to 5 μg of poly(A)$^+$ RNA isolated from female spikelets and hybridized overnight at 45°C. The unhybridized RNA was digested for 1 hr at 30°C with RNase A (40 μg/ml) and RNase T_1 (8 units/ml). The primer-extended and RNase protection products were each separated in an 8% acrylamide sequencing gel.

RESULTS

Northern Blot Analysis. The *R-nj* gene was cloned by a transposon tagging strategy and used in Southern blot analysis to demonstrate that different *R* genes are homologous (19). As a first step in understanding what an *R* gene encodes and to determine whether transcripts from diverse *R* genes were similar, Northern blot analysis was performed. This analysis was also used to identify the most abundant *R* transcript to aid in the subsequent isolation of an *R* cDNA.

We presumed that organs displaying intense anthocyanin pigmentation would have the highest levels of an *R* transcript. Therefore, poly(A)$^+$ RNA was isolated from male spikelets (containing *P*), female spikelets (containing *Lc*), or kernels (containing *S*) and size-fractionated in a 1.2% agarose/formaldehyde gel. The RNA blot was probed with a 0.7-kb *Bgl* II–*Hin*cII fragment isolated from the *R-nj* genomic clone (19). The results (Fig. 1) revealed that (*i*) the predominant transcript encoded by all three *R* genes is ≈2.5 kb long and (*ii*) the highest level of transcript is found in female spikelets containing the *Lc* gene (lane 2).

Genomic and cDNA Cloning and Sequence Analysis. We were unable to detect an *R-nj* transcript (data not shown); therefore we decided to focus our analysis on the *Lc* gene, since it encodes a relatively abundant mRNA (Fig. 1). Southern blot analysis of genomic DNA isolated from isogenic lines with and without the *Lc* gene identified a 3.7-kb *Hin*dIII fragment containing *Lc* (Fig. 2, compare lanes 1 and 2). This *Hin*dIII fragment was cloned and used to probe a cDNA library prepared from female spikelet poly(A)$^+$ RNA (Fig. 1, lane 2). One cDNA clone hybridized to the 3.7-kb *Lc* genomic fragment and was purified, subcloned, and sequenced.

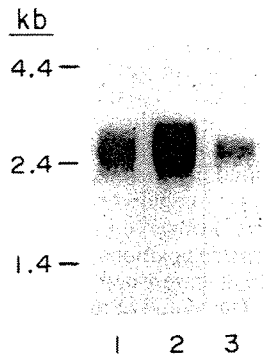


FIG. 1. RNA blot analysis of strains containing different *R* genes. Samples (5 μg) of poly(A)$^+$ RNA isolated from male spikelets (*P* genotype) (lane 1), female spikelets 10 dap exposed 2 days to light (*Lc* genotype) (lane 2), or kernels 35 dap (*S* genotype) (lane 3) were electrophoresed, transferred to Magnagraph membranes, and probed with pR-nj:1 (19).

When the sequence of the 3.7-kb *Hin*dIII genomic fragment was compared with the cDNA sequence, it was determined that the 3.7-kb genomic fragment contained only two small exons (228 and 159 base pairs long), which were identical to the 5′ end of the cDNA. To identify the genomic fragment that encoded the rest of the transcription unit, the Southern blot displayed in Fig. 2 (lanes 1 and 2) was stripped and probed with the 3′ end of the cDNA (positions 880–1772). This Southern blot revealed the presence of a 6-kb *Hin*dIII fragment specific to genomic DNA containing *Lc* (Fig. 2, compare lanes 3 and 4). This 6-kb fragment was isolated from a genomic library containing size-fractionated *Hin*dIII-digested genomic DNA cloned into λZAP.

Characterization of the *Lc* Transcription Unit. The sequence of the cDNA clone was identical to sequences contained on the two genomic fragments, confirming that the cDNA was derived from *Lc*. The transcription unit was found

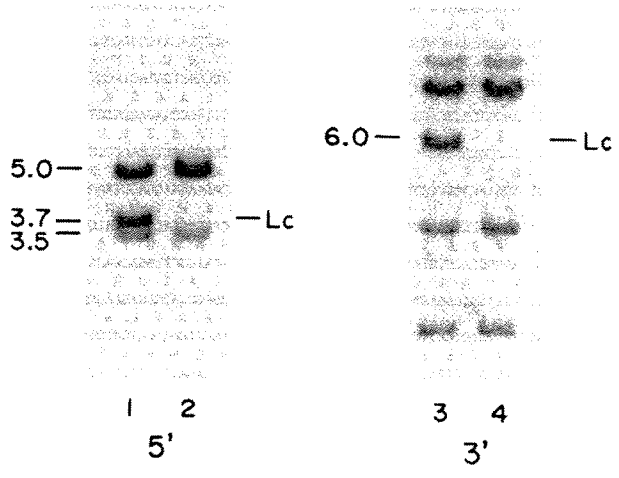


FIG. 2. Southern blot analyses of isogenic lines with and without the *Lc* gene. Samples (10 μg) of genomic DNA isolated from young leaf tissue containing the *Lc* gene (lanes 1 and 3) or lacking the Lc gene (lanes 2 and 4) were digested with *Hin*dIII, electrophoresed, and transferred to nitrocellulose. Lanes 1 and 2 were probed with a 5′ probe, pR-nj:1 (19). Lanes 3 and 4 were probed with a 3′ *Lc* cDNA probe (nucleotides 880–1772).

to span ≈7 kb, with introns at positions 228/229, 387/388, 648/649, 745/746, 760/761, 817/818, 1566/1567, and 1996/1997 (Fig. 3). All the introns contained the consensus splice junction sequence (5′-GT . . . AG-3′) (data not shown).

Primer-extension was used to determine the start of transcription (30). A 20-base oligonucleotide was hybridized to poly(A)$^+$ RNA isolated from plants with and without *Lc* (Fig. 4*A*, lanes 1 and 2) and extended with reverse transcriptase. Two major bands were observed (Fig. 4*A*, lane 1) and were used to define nucleotide positions 1 and 3 (Fig. 3). An RNase protection experiment confirmed this as the start of transcription (Fig. 4*B*). The 5′ end of the *Lc* cDNA starts 20 base pairs from the transcription start site. A putative TATA box (TATATATA) is located upstream of the start of transcription at positions −31 to −24 (Fig. 3).

Features of the *Lc*-Encoded Protein. The cDNA sequence contains a 610-amino acid open reading frame beginning with an ATG at nucleotide position 236 and ending with a stop codon at nucleotide position 2066. Although the open reading frame begins with the fourth ATG from the 5′ end, Kozak (31) reported that 5–10% of eukaryotic mRNAs contain upstream AUG codons. The putative protein encoded by the 610-amino acid open reading frame has features similar to those of transcriptional activators (32). The activating regions of known transcriptional activator proteins contain stretches of amino acids with a significant net negative charge (32). These proteins also contain basic domains believed to be involved in DNA binding. The putative Lc protein also contains a large acidic domain and a smaller basic domain (amino acids 188–318 and 415–508, respectively, Fig. 5*A*). The Lc acidic domain contains 33 acidic and 8 basic amino acids for a net charge of −25.

The predicted protein sequence of Lc was analyzed for homology to protein sequences produced by translating the GenBank data base (release 55) in all three reading frames. Part of the basic domain of Lc (amino acids 420–462) was found to have homology to the *myc*-homology region shared by the nuclear *myc* oncogenes (L-*myc* shown, ref. 33); a cDNA encoding a myoblast determination factor (MyoD1, ref. 34); several genes of *Drosophila*, including daughterless (*da*, ref. 35), genes of the *achaete–scute* complex (gene *T3* of *AC-S* shown, ref. 36), and twist (37); and a human B-cell cDNA, E12 (38) (Fig. 5*B*).

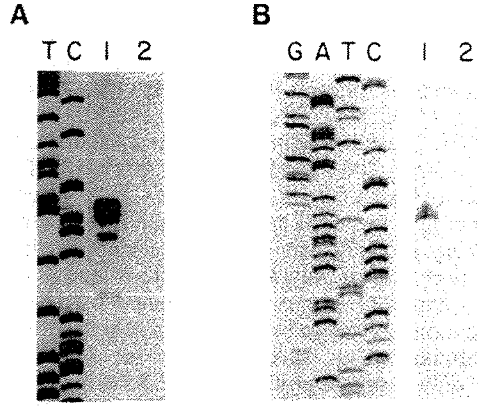


FIG. 4. Primer extension and RNase protection analyses. (*A*) A 20-base oligonucleotide was hybridized to poly(A)$^+$ RNA from female spikelets with or without the *Lc* gene (lanes 1 and 2, respectively) and extended with reverse transcriptase. (*B*) A 1.6-kb *Hin*dIII–*Sph* I *Lc* genomic fragment was transcribed *in vitro* with SP6 polymerase. The labeled RNA was hybridized to 5 μg of poly(A)$^+$ RNA from female spikelets with or without *Lc* (lanes 1 and 2, respectively) and digested with RNase A and RNase T_1. The primer-extended and RNase protection products were separated in an 8% acrylamide sequencing gel along with products of sequencing reactions (lanes G, A, T, and C) using the genomic DNA fragment and pUC119 for size standards, respectively (27).

RNA Analysis of *R*, *C2*, and *A1* mRNA. Previous biochemical studies indicated that an *R* gene is required for *A1*-, *C2*- and *Bz*-encoded enzymatic activities (16–18). Because the sequence data indicated that Lc may be a DNA-binding transcriptional activator, we investigated the possibility that *R* may influence *C2* and *A1* transcript levels. Poly(A)$^+$ RNAs isolated from kernels (35 dap) containing or lacking the *S* gene were analyzed by RNA blot hybridizations (Fig. 6). It was observed that *S*, *C2*, and *A1* transcripts were present only in

```
-140 cccaaggttcgtggcatatctgtaggcatctaccccgtcttcgtcgtccgctcctcactagctaccaagaggtcgccattattgccaacatagagtgtacgtggatgtctatatatatgcctacttgcacccatatggca
                         *
   1 TAGGCGTTCGATCCCCTTAGCGCGGAGGAGAGCTCCTCCGGTTCTTCTCTACCCTTCGCATGGAAGTTCTTGCATTGCTTCGTTGCTTCTCTAGTTTCTTCCTTCTACGTCTTTCCAGCATACGCATGCCCCTCGTCCGC
                                                                                                  M  A  L  S  A  S  R  V  Q  Q  A  E  E  L  L
 141 CGGTTCACGAGGCATCGTCTGATGATCAGTAGATAATAAGCAATATAATACTGATCTAGAATCGAGTTGTTGTACTCTTCGCAGATAGGCGCGTGATGGCGCTTTCAGCTTCCCGAGTTCAGCAGGCGGAAGAACTGCTG
      Q  R  P  A  E  R  Q  L  H  R  S  Q  L  A  A  A  A  R  S  I  N  W  S  Y  A  L  F  W  S  I  S  D  T  Q  P  G  V  L  T  W  T  D  G  F  Y  N  G
 281 CAACGACCTGCTGAGAGGCAGCTGATGAGGAGCCAGCTTGCTGCAGCCGCCAGGAGCATCAACTGGAGCTACGCCCTCTTCTGGTCCATTTCAGACACTCAACCAGGGGTGCTGACGTGGACGGACGGGTTCTACAACGG
      E  V  K  T  R  K  I  S  N  S  V  E  L  T  S  D  Q  L  V  H  Q  R  S  D  Q  L  R  E  L  Y  E  A  L  L  S  G  E  G  D  R  R  A  A  P  A  R
 421 CGAGGTGAAGACGCGGAAGATCTCCAACTCCGTGGAGCTGACATCCGACCAGCTCGTCATGCAGAGGAGCGACCAGCTCCGGGAGCTCTACGAGGCCCTCCTGTCGGGCGAGGGCGACCGCCGCGCTGCGCCTGCGCGGC
      P  A  G  S  L  S  P  E  D  L  G  D  T  E  W  Y  Y  V  V  S  H  T  Y  A  F  R  P  G  Q  G  L  P  G  R  S  F  A  S  D  E  H  V  W  L  C  N  A
 561 CGGCCGGCTCTCTGTCGCCGGAGGACCTCGGCGACACCGAGTGGTACTACGTGGTCTCCATGACCTACGCCTTCCGGCCAGGCCAAGGGTTGCCCGGCAGGAGTTTCGCGAGCGACGAGCATGTCTGGCTGTGCAACGCG
      H  L  A  G  S  K  A  F  P  R  A  L  L  A  K  S  A  S  I  Q  S  I  L  C  I  P  V  H  G  G  V  L  E  L  G  T  T  D  T  V  P  E  A  P  D  L  V
 701 CACCTCGCCGGCAGCAAAGCCTTCCCCCGCGCGCTCCTGGCCAAGAGCGCGTCCATTCAGTCAATCCTCTGCATCCCGGTTATGGGCGGCGTGCTTGAGCTTGGTACAACTGACACGGTGCCGGAGGCCCCGGACTTGGT
      S  R  A  T  A  A  F  W  E  P  Q  C  P  S  S  S  P  S  G  R  A  N  E  T  G  E  A  A  A  D  D  G  T  F  A  F  E  E  L  D  H  N  N  G  H  D
 841 CAGCCGAGCAACCGCGGCTTTCTGGGAGCCGCAGTGCCCGAGCTCCAGCCCGTCAGGACGAGCAAACGAGACCGGCGAGGCCGCAGCAGACGACGGCACGTTTGCGTTCGAGGAACTCGACCACAATAATGGCATGGACG
      D  I  E  A  N  T  A  A  G  G  H  G  Q  E  E  E  L  R  L  R  E  A  E  A  L  S  D  D  A  S  L  E  N  I  T  K  E  I  E  E  F  Y  S  L  C  D  E
 981 ACATAGAGGCGATGACCGCCGCCGGGGGACACGGGCAGGAGGAGGAGCTAAGACTAAGAGAAGCCGAGGCCCTGTCAGACGACGCAAGCCTGGAGCACATCACCAAGGAGATCGAGGAGTTCTACAGCCTCTGCGACGAA
      H  D  L  Q  A  L  P  L  P  L  E  D  G  W  T  V  D  A  S  N  F  E  V  P  C  S  S  P  Q  P  A  P  P  P  V  D  R  A  T  A  N  V  A  A  D  A  S
1121 ATGGACCTGCAGGCGCTACCACTACCGCTAGAGGACGGCTGGACCGTGGACGCGTCCAATTTCGAGGTCCCCTGCTCTTCCCCGCAGCCAGCGCCGCCTCCGGTGGACAGGGCTACCGCTAACGTCGCCGCCGACGCCTC
      R  A  P  V  Y  G  S  R  A  T  S  F  H  A  W  T  R  S  S  Q  Q  S  S  C  S  D  D  A  A  P  A  A  V  V  P  A  I  E  E  P  Q  R  L  L  K  K
1261 AAGGGCACCCGTCTACGGCTCTCGCGCGACGAGTTTCATGGCTTGGACGAGGTCCTCGCAGCAGTCGTCGTGCTCCGACGACGCGGCGCCCGCAGCAGTAGTGCCGGCCATCGAGGAGCCGCAGAGATTGCTGAAGAAAG
      V  V  A  G  G  G  A  W  E  S  C  G  G  A  T  G  A  A  Q  E  H  S  G  T  G  T  K  N  H  V  H  S  E  R  K  R  R  E  K  L  N  E  H  F  L  V  L
1401 TGGTGGCCGGCGGCGGTGCTTGGGAGAGCTGTGGCGGCGCGACGGGAGCAGCACAGGAAATGAGTGGCACTGGCACCAAGAACCACGTCATGTCGGAGCGAAAGCGACGAGAGAAGCTCAACGAGATGTTCCTCGTCCTC
      K  S  L  L  P  S  I  H  R  V  N  K  A  S  I  L  A  E  T  I  A  Y  L  K  E  L  Q  R  R  V  Q  E  L  E  S  S  R  E  P  A  S  R  P  S  E  T  T
1541 AAGTCACTGCTTCCGTCCATTCACAGGGTGAACAAAGCGTCGATCCTCGCCGAAACGATAGCCTACCTCAAGGAGCTTCAGAGAAGGGTGCAAGAGCTGGAGTCCAGTAGGGAACCTGCGTCGCGCCCATCCGAAACGAC
      T  R  L  I  T  R  P  S  R  G  N  N  E  S  V  R  K  E  V  C  A  G  S  K  R  K  S  P  E  L  G  R  D  D  V  E  R  P  P  V  L  T  H  D  A  G
1681 GACAAGGCTAATAACAAGGCCCTCCCGTGGCAATAATGAGAGTGTGAGGAAGGAGGTCTGCGCGGGCTCCAAGAGGAAGAGCCCAGAGCTCGGCAGAGACGACGTGGAGCGCCCCCCGGTCCTCACCATGGACGCCGGCA
      T  S  N  V  T  V  T  V  S  D  K  D  V  L  L  E  V  Q  C  R  W  E  E  L  L  H  T  R  V  F  D  A  I  K  S  L  H  L  D  V  L  S  V  Q  A  S  A
1821 CCAGCAACGTCACCGTCACCGTCTCGGACAAGGACGTGCTCCTGGAGGTGCAGTGCCGGTGGGAGGAGCTCCTGATGACGCGAGTGTTCGACGCCATCAAGAGCCTCCATTTGGACGTCCTCTCGGTTCAGGCTTCAGCG
      P  D  G  F  H  G  L  K  I  R  A  Q  F  A  G  S  G  A  V  V  P  W  H  I  S  E  A  L  R  K  A  I  G  K  R  -
1961 CCAGATGGCTTCATGGGGCTTAAGATACGAGCTCAGTTTGCTGGCTCCGGTGCCGTCGTGCCCTGGATGATCAGCGAGGCTCTTCGCAAAGCTATAGGGAAGCGGTGAAGGGGCAGCTGGAAATTTGGACATCGACGGGC

2101 ATGGAAGGCTTCATGGGATCGAAGCAAAGATGTATTTCTTGTTTCTTTAGATAACAGACATGAATCGGACCTTTATATCAACAATTATATGGGCATGAATACTTAAGACTCCAGCCCTTAACACGTAGAAACTCAAAAAA

2241 GAAGAGAGGAAGCTAAAGACTAAGCGTAAGGTATATTTGGAAGTAAATTATTTTTATAGTTTCTAAGCAATCTCATGGTTTATAGGAATACTAGAGTGTTTATGGCATAAGGTGTTTGGTTGCATTCATAAAACCTATAT

2381 TTTCAAAGTCATAGCATTCTAGATACCATGATATTTTTGTAATATTGGAAACTACACTCCAACGCAAAGTTTTTATGACATGGCT
```

FIG. 3. Nucleotide sequence of *Lc* cDNA and 5′ flanking genomic sequence. Uppercase letters indicate the extent of the *Lc* transcript; the start site is defined as position 1. The amino acids derived from translation of this sequence are shown in one-letter code above the corresponding codons. The beginning of the cDNA is marked with an asterisk. The presumed TATA box and the intron positions are underlined.

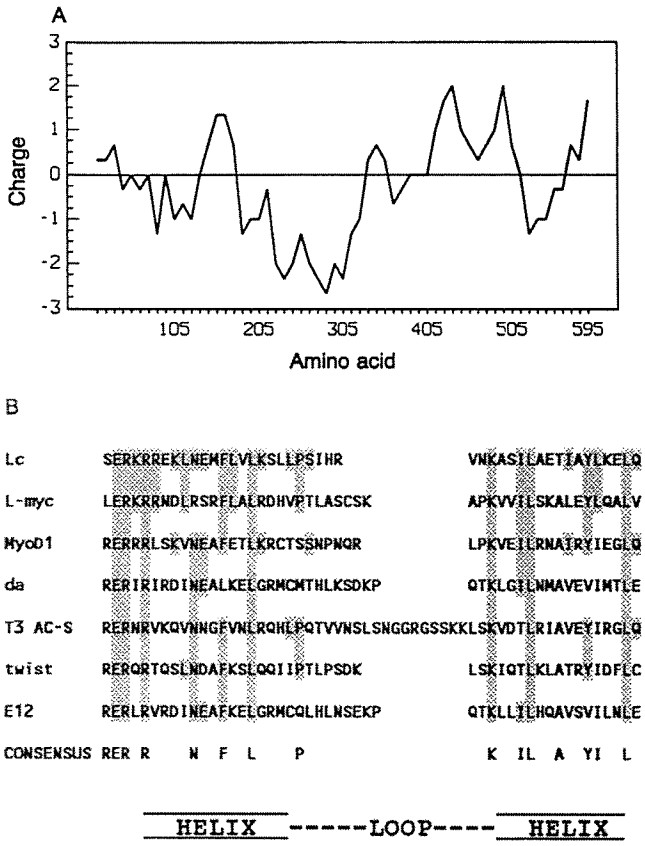


FIG. 5. Characteristics of the putative Lc protein. (*A*) The average of net charges of the putative Lc protein over successive 30 amino acids, measured at 10-amino acid intervals. (*B*) Amino acid comparison of the helix–loop–helix *myc*-homology region encoded by *Lc*, amino acids 420–462; *L-myc*, 297–334 (33); MyoD1 cDNA, 118–161 (34); *da*, 563–608 (35); *T3* of *AS-C*, 62–146 (36); twist, 365–409 (37); and E12, 344–390 (38). Amino acid identities that include Lc have been shaded and a consensus sequence containing amino acids occurring in more than half of the proteins shown is presented.

the RNA from the *S* gene-containing kernels. The cloned maize actin gene (pMAc1; ref. 41) was used as a control and

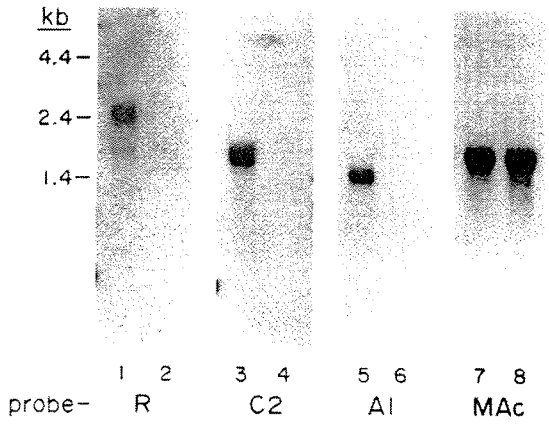


FIG. 6. Northern blot analysis of RNA from strains containing or lacking functional *R* genes. Poly(A)$^+$ RNA from kernels (35 dap) containing (lanes 1, 3, 5, and 7) or lacking (lanes 2, 4, 6, and 8) the *S* gene were analyzed by RNA blot hybridizations. The membrane was probed with *Lc* cDNA (positions 880–1772; lanes 1 and 2), chalcone synthase (pc46E, ref. 39; lanes 3 and 4), dihydroquercetin reductase (pALC2, ref. 40; lanes 5 and 6), and maize actin (pMAc1, ref. 41; lanes 7 and 8) probes.

was present in equal amounts in both lanes (Fig. 6, lanes 7 and 8). Similar experiments using RNA from male spikelets containing or lacking the *R* gene *P* indicated that *P* must be present in order to detect *C2* and *A1* transcripts in anthers (data not shown).

DISCUSSION

In this study we present data that support previous experiments (16–18) showing that the *R* gene product has a regulatory role in the maize anthocyanin biosynthetic pathway. The data can be summarized as follows: (*i*) the Lc gene encodes a 2.5-kb mRNA that can be translated into a 610-amino acid protein characterized by large acidic and basic domains, (*ii*) part of the basic domain contains a consensus region present in other putative DNA-binding regulatory proteins, and (*iii*) strains lacking functional *R* genes do not produce the *R*-encoded 2.5-kb transcript or transcripts encoded by *C2* or *A1*, two genes that *R* is believed to regulate.

The predicted R protein shares features with many eukaryotic regulatory proteins (32). The activating regions of transcriptional activating proteins are characterized by acidic amino acids (32). The R protein described in this study contains a very large acidic domain with a net charge of −25. Basic domains of regulatory proteins have been associated with DNA binding activity (32). Not only does the R protein have a basic domain, but this region contains the *myc*-homology region shared by several other DNA-binding proteins (38). Murre *et al.* (38) proposed that this conserved domain could encode a helix–loop–helix motif that is required for DNA binding and dimerization. With the cloning of the two regulatory genes *R* and *C1* and several of the genes they regulate, the tools are now available to address questions regarding both DNA–protein and protein–protein interactions.

Our data, taken together with previous genetic studies, support the idea that all or most *R* genes encode functionally equivalent proteins. Genetic analysis of strains containing different *R* genes has revealed that recombination can occur between them.¶ In addition, Southern blot analysis demonstrated that a genomic fragment isolated from the *R-nj* gene was homologous with the *R* genes *S* and *P* (19). Finally, in this study we demonstrate that three *R* genes, *S*, *P*, and *Lc*, encode homologous transcripts 2.5 kb in length.

Is the putative product of this mRNA, the 610-amino acid protein, the only product of *R*? We believe that the answer is yes for the following reason. The approximate limits of another *R* gene, termed *R-sc*, have been defined by transposon mutagenesis with the *Ds* element.¶ Forty insertions that affect *R-sc* expression delimit a 9-kb region of DNA. A comparison of the *Lc* sequence with the positions of the *R-sc* *Ds* insertion reveals that the *Ds* elements are either within the RNA coding region or in 5′ flanking sequences of the 7-kb transcription unit.

If we assume that all or most *R* genes contain a single transcription unit that encodes the R protein, what then is the basis for the diverse expression displayed by members of the *R* gene family? We propose that the diverse patterns of anthocyanin pigmentation conditioned by different *R* genes reflect differences in the *R* gene promoters rather than their gene products. This model predicts that it is the presence or absence of the R protein that determines when and where a

¶Alleman, M., Dellaporta, S. & Kermicle, J., International Symposium on Plant Transposable Elements, August 22–26, 1987, Madison, WI, p. 41 (abstr.).

plant will be pigmented when all other genes in the pathway are functional.

We are indebted to Dr. Jerry Kermicle for sharing with us his knowledge of the *R* genes, as well as providing us with the seed material necessary for this project. We thank Dr. Udo Wienand for providing the *A1* and *C1* clones and Dr. Rich Meagher for providing the maize actin clone. We also thank Andy Tull for excellent greenhouse support. We are grateful to Drs. Glenn Galau, Teresa Gruber, Alan Jaworski, and Clifford Weil for critical reading of the manuscript. This study was supported by a U.S. Department of Energy grant under Contract DE-FG09-86ER13621 to S.R.W. and by a U.S. Department of Energy grant to S.L.D.

1. Coe, E. H., Jr., Neuffer, M. G. & Hoisington, D. A. (1988) in *Corn and Corn Improvement*, eds. Sprague, G. F. & Dudley, J. W. (Am. Soc. Agronomy, Madison, WI), pp. 81–258.
2. Wienand, U., Weydermann, U., Niesbach-Kloesgen, U., Peterson, P. A. & Saedler, H. (1986) *Mol. Gen. Genet.* **203**, 202–207.
3. O'Reilly, C., Shepherd, N. S., Pereira, A., Schwarz-Sommer, Z., Bertram, I., Robertson, D. S., Peterson, P. A. & Saedler, H. (1985) *EMBO J.* **4**, 877–882.
4. Fedoroff, N. V., Furtek, D. B. & Nelson, E. O. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3825–3829.
5. Paz-Ares, J., Wienand, U., Peterson, P. & Saedler, H. (1986) *EMBO J.* **5**, 829–833.
6. Cone, K. C., Burr, F. A. & Burr, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9631–9635.
7. Paz-Ares, J., Ghosal, D., Wienand, U., Peterson, P. A. & Saedler, H. (1987) *EMBO J.* **6**, 3553–3558.
8. Styles, E. D., Ceska, O. & Seah, K.-T. (1973) *Can. J. Genet. Cytol.* **15**, 59–72.
9. van der Walt, W. (1967) Ph.D. Dissertation (Univ. Wisconsin, Madison).
10. van der Walt, W. J. & Brink, R. A. (1969) *Genetics* **61**, 677–695.
11. Fogel, S. (1946) Ph.D. Dissertation (Univ. Missouri, Columbia).
12. Stadler, L. J. (1948) *Am. Nat.* **82**, 289–314.
13. Webber, H. J. (1906) *Proc. Am. Breeders Assoc.* **2**, 73–83.
14. Stadler, L. J. & Neuffer, M. G. (1953) *Science* **117**, 471–472.
15. Dooner, H. K. & Kermicle, J. L. (1976) *Genetics* **82**, 309–322.
16. Reddy, A. R., Britsch, L., Salamini, F., Saedler, H. & Rohde, W. (1987) *Plant Sci.* **52**, 7–13.
17. Dooner, H. K. (1983) *Mol. Gen. Genet.* **189**, 136–141.
18. Dooner, H. K. & Nelson, O. (1979) *Genetics* **91**, 309–315.
19. Dellaporta, S. L., Greenblatt, I., Kermicle, J., Hicks, J. B. & Wessler, S. R. (1988) in *Chromosome Structure and Function: Impact of New Concepts*, 18th Stadler Genetics Symposium, eds. Gustafson, J. P. & Appels, R. (Plenum, New York), pp. 263–282.
20. Wessler, S. R., Baran, G., Varagona, M. & Dellaporta, S. L. (1986) *EMBO J.* **5**, 2427–2432.
21. Fedoroff, N., Mauvais, J. & Chaleff, D. (1983) *J. Mol. Appl. Gen.* **2**, 11–29.
22. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
23. Hu, N. T. & Messing, J. (1982) *Gene* **17**, 271–277.
24. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
25. Loenen, W. A. M. & Blattner, F. R. (1983) *Gene* **26**, 171–179.
26. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
27. Vieira, J. & Messing, J. (1987) *Methods Enzymol.* **153**, 3–11.
28. Dale, R. M. K., McClure, B. A. & Houchins, J. P. (1985) *Plasmid* **13**, 31–40.
29. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
30. Dunsmuir, P., Bond, K., Lee, K., Gidoni, D. & Townsend, J. (1988) in *Plant Molecular Biology Manual*, eds. Gelvin, S. B. & Schilperoort, R. A. (Kulwer, Hingham, MA), pp. 1–17.
31. Kozak, M. (1986) *Cell* **47**, 481–483.
32. Ptashne, M. (1988) *Nature (London)* **335**, 683–689.
33. DePinho, R. A., Hatton, K. S., Tesfaye, A., Yancopoulos, G. D. & Alt, F. W. (1987) *Genes Dev.* **1**, 1311–1326.
34. Davis, R. L., Weintraub, H. & Lasser, A. B. (1987) *Cell* **51**, 987–1000.
35. Caudy, M., Vassin, H., Brand, M., Tuma, R., Jan, L. Y. & Jan, Y. N. (1988) *Cell* **55**, 1061–1067.
36. Villares, R. & Cabrera, C. V. (1987) *Cell* **50**, 415–424.
37. Thisse, B., Stoetzel, C., Gorostiza-Thisse, C. & Perrin-Schmitt, F. (1988) *EMBO J.* **7**, 2175–2183.
38. Murre, C., McCaw, P. S. & Baltimore, D. (1989) *Cell* **56**, 777–783.
39. Niesbach-Klosgen, U., Barzen, E., Bernhardt, J., Rohde, W., Schwarz-Sommer, Z., Reif, H. J., Wienand, U. & Saedler, H. (1987) *J. Mol. Evol.* **26**, 213–225.
40. Schwarz-Sommer, Z., Shepherd, N., Tacke, E., Gierl, A., Rohde, W., Leclercq, L., Mattes, M., Berndtgen, R., Peterson, P. A. & Saedler, H. (1987) *EMBO J.* **6**, 287–294.
41. Shah, D. M., Hightower, R. C. & Meagher, R. B. (1983) *J. Mol. Appl. Genet.* **2**, 111–126.

B Fleissner

MOLECULAR CLONING OF THE MAIZE *R-nj* ALLELE BY TRANSPOSON TAGGING WITH *Ac*

Stephen L. Dellaporta*, Irwin Greenblatt+, Jerry L. Kermicle++, James B. Hicks# and Susan R. Wessler$^{\delta}$

*Yale University
Department of Biology
New Haven, CT 06511

+Molecular and Cellular Biology
University of Connecticut
Storrs, CT 06268

++Laboratory of Genetics
University of Wisconsin
Madison, WI 53706

#Scripps Clinic and Research Foundation
Department of Molecular Biology
La Jolla, CA 92037

$^{\delta}$Department of Botany
University of Georgia
Athens, GA 30602

ABSTRACT

The *R* locus of maize is one of several genes that condition the red and purple anthocyanin pigments throughout the body of the plant and seed. Many alleles of the *R* locus have been described, each determining a different pattern of pigment distribution in the plant. Genetic studies have lead to the proposal that *R* is a complex

locus containing several similar but functionally distinct modules that regulate anthocyanin expression in specific organs or tissues. We report on the isolation of a segment of DNA from the *R* locus using the controlling element *Activator* (*Ac*)as a transposon tag. Molecular evidence is presented confirming the complex organization of this regulatory locus. The identification of cross-hybridizing restriction fragments corresponding to the tissue-specific *R* components suggests that these elements, while functionally divergent in tissue-specificity, have retained significant DNA homology and may have evolved from a common ancesteral component at *R*.

INTRODUCTION

The *R* gene of maize impinges on anthocyanin pigmentation in at least two ways. First, it governs the expression of unlinked genes encoding enzymes in the flavonoid biosynthetic pathway. An active *R* gene is required for normal activity of the first and a terminal step of flavonoid biosynthesis (products of the *C2* and *Bronze* genes) and, presumably for other steps in the pathway (coded by *A*, *A2*, *C*, and *Bronze-2*) (Dooner, 1983). Second, various alleles of *R* confer diverse patterns of pigment distribution among seed and plant parts. Thus, the *R* gene may regulate pigment deposition by controlling the expression of the flavonoid biosynthetic gene pathway in a tissue-specific manner. Alternatively, *R* may confer tissue-specific pigment distribution patterns by producing a product(s) in specific cells that catalyzes an early enzymatic step in flavonoid biosynthesis.

R-allelic diversity is attributable to variation in the number and kind of genetic units which L.J. Stadler (1951) termed "*R*-genic elements". Such elements occur at the *R* locus of different maize races either singly or as complexes of two or more members. Genetic analysis of one complex, designated as *R-r:standard*, reveals one element which confers anthocyanin formation to the aleurone layer of the kernel endosperm and another which confers coloration on various seedling and mature plant parts (Dooner and Kermicle, 1971). These elements, called plant or (P), and seed or (S), components, behave genetically as separate genes within the *R* complex. A schematic of this genetic model is shown in Fig. 1A. The *R-navajo* (*R-nj*) allele, utilized in the present study for molecular cloning, has not given evidence of complexity in similar tests although it pigments the crown portion of the aleurone in the kernel and various seedling parts as well as the anthers and silks of mature plants (Kermicle 1970, unpublished data).

Information concerning the function of the *R* gene product is not available. Therefore, we have begun a molecular study of this locus by cloning a portion of the *R* complex. Because the *R* gene product has not been characterized, an indirect approach to identifying an *R* molecular clone was necessary. Following the molecular approach first described by Fedoroff et al (1984) to clone the *bronze* locus, we made use of a cloned mobile sequence (Fedoroff et al; 1983), *Activator* (*Ac*), as a "transposon tag" for the *R* DNA. *Ac* transposes preferentially to nearby sites in the

chromosome (Van Schaik and Brink, 1979; Greenblatt and Brink, 1962; Greenblatt, 1984), serving as a localized mutagen. A reciprocal translocation was used to bring an *Ac* element located at the *P* locus on chromosome-1 and the target allele *R-nj*, in chromosome-10 together in *cis* arrangement. An unstable allele of *R-nj*, designated *r-nj:m1*, was located from this stock and shown to contain an active *Ac* at *R-nj* by genetic and molecular analysis using an *Ac* sequence as a hybridization probe. The *Ac* containing fragment was cloned and an adjacent unique genomic fragment was used as a hybridization probe on Southern blots to correlate physical structure with the organization of the *R-r:standard* allele. Three restriction fragments with sequence homology to the *R-nj* DNA probe were identified two of which can be correlated with (P) and (S) function. The third homologous fragment indicates the presence of a cryptic component of the *R-r* allele.

MATERIALS AND METHODS

R-alleles

R-navajo (R-nj): Isolated from the Cudu strain, this stable allele is distinguished most obviously by a solid patch of aleurone pigment in the crown region of the kernel. Other regions of the aleurone are pigmented irregularly. The embryos, coeloptile and roots of seedlings, as well as silks and anthers of mature plants pigment with variable intensities.

R-r:standard and its derivatives: *R-r:standard* confers strong uniform pigmentation to the aleurone layer of kernels, to various seedling parts and to anthers of mature plants. Plant and seed effects are determined by separate genetic elements, (P) and (S), which are organized as members of a direct duplication. Unequal crossing over results in loss of the duplication, yielding products carrying only (P), designated as *r-r*, or (S), designated *R-g*. The particular such derivatives used here were *r-r:n46* and *R-g:1*.

Mutation of (S) to (s) in *R-r:standard* yielded a (P) (s) product designated as *r-r:n35*. Such noncrossover *r-r* mutants retain the duplication of *R-r:standard* as evidenced by the fact that they retain two doses of *Inhibitor of striate* (*Isr*), a second gene encompassed within the duplicated segment (Kermicle and Axtell, 1981), and recombine with (S) derivatives to reconstitute the *R-r* phenotype (Dooner and Kermicle, 1974).

r-g: This test allele pigments neither the seed nor plant parts.

Isolating *Ac*-induced *R* mutations

Detection of putative insertions of *Ac* into *R-nj* were made on the basis of the presence of variegated aleurone color in the region of the *R-nj* expression - only on

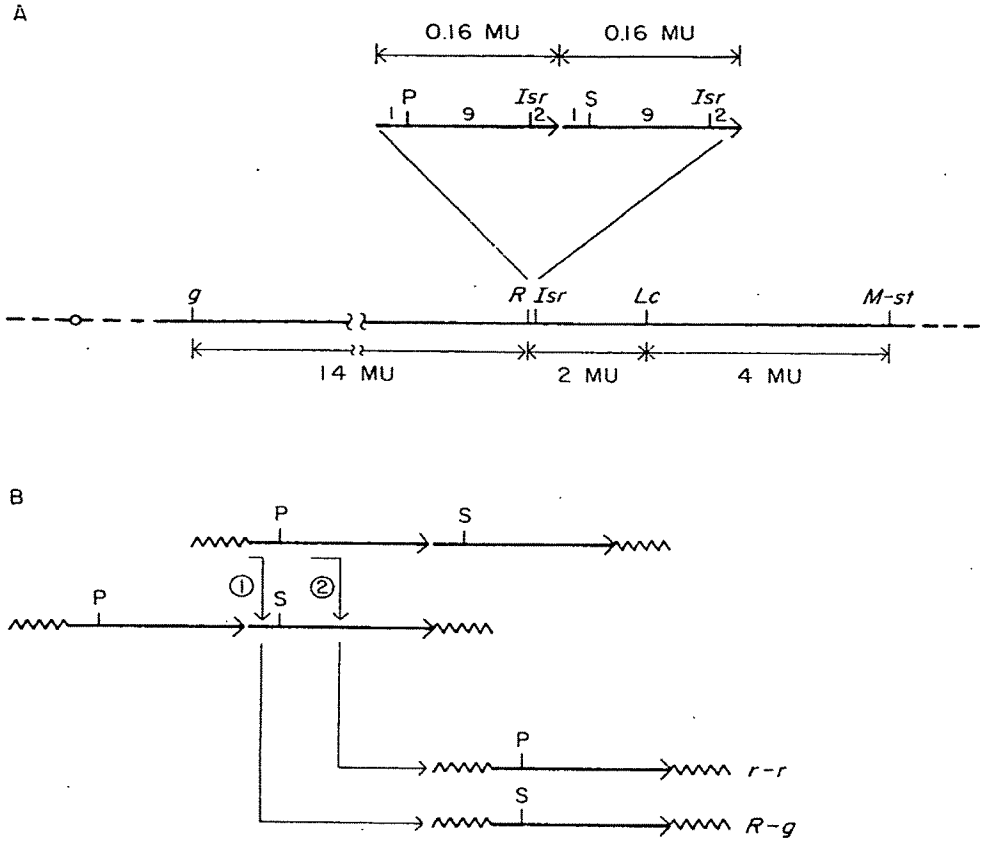


Fig. 1. Genetic Organization of the *R* Locus.

A. Genetic map of the long arm of chromosome-10 of maize showing the genetic map of the *R* locus and flanking markers: *g: golden; Isr: inhibitor of strlate; Lc: leaf color component; M-st: modifier of stippled.* The region that includes *R* is expanded to show the genetic fine structure relationship of the (P) and (S) components of the *R-r:standard* allele. Each component is carried on a chromosome duplicated segment (arrow) and includes the linked marker *Isr* as part of the duplication. The position of the *R* component indicated above the arrow is based on the relative numbers of recombinants observed in experiments of *R-r* fractionation and resynthesis (see Dooner and Kermicle, 1974, 1976). The position of *Isr* within the duplicated segment was determined relative to a distinguishable allele carried in *R-stippled* (Kermicle and Axtell, 1981, and Kermicle, unpublished).

B. Genetic model of unequal crossing over in *R-r* homozygotes. Recombination in region 1 or 2 (arrows) will yield *R-g* (S) or *r-r* (P) derivatives, respectively.

the crown portion of the kernel (Fig. 1). Kernels were planted and tested if they had a colorless aleurone with colored spots occurring and also if they had colored aleurone with colorless spots occurring. All such kernels were planted and pollinated with a *r-g* stock. It was from this population that the *R-nj* mutable analyzed here (*r-nj:m1)* was recovered. The *r-nj:m1* mutation was propagated for several generations by backcrossing into a *r-g* W22 genetic background.

DNA preparation, restriction and genomic blot analysis

Genomic DNA was isolated from mature leaf tissue by a previously described method (Shure et al., 1983). Genomic DNA was digested with a three-fold excess units of restriction enzymes according to the manufacturer's recommendation (New England Biolabs or Bethesda Research Laboratories). DNA samples were electrophoresed through 0.8% agarose (Sigma) gels, and were transferred to nitrocellulose (Schleicher and Schuell) according to the method of Southern (1975). Plasmid DNA was isolated by the method of Holmes and Quigley (1981), the plasmid insert was purified by electroelution (Dretzen et al., 1981) and nick translated as described by Rigby et al., (1977). DNA polymerase and DNAase I for nick translation was obtained from International Biotechnology Inc., and ^{32}P-α-NTP at a specific activity of 800 Ci/mM was supplied by New England Nuclear.

Filters were prehybridized for 2 hrs at 65°C in 10% dextran sulfate (Pharmacia), 6X SCP (1X SCP = 100 mM NaCl, 30 mM Na_2HPO_4, and 1 mM EDTA, pH 6.5), 2% sarcosine and 500 ug/ml heparin (Type II Sigma). Hybridization was performed for 12 hrs at 65°C in the above buffer containing denatured salmon sperm DNA (final concentration 100 ug/ml) and denatured, nick translated probe DNA (finale probe concentration = 10 ng/ml). The filters were washed three times in 2X SCP, 1% SDS at 65°C for 30 minutes and then in 0.1X SCP, 0.1% SDS at 65°C for 15 minutes.

Genomic Library Construction and Screening

Maize DNA libraries were constructed with 2 ug of Bam HI-digested genomic DNA and 5 ug of λ EMBL4 arms and co-ligated at 4°C for 24 hr in a 15 ul reaction volume containing 100 units of T_4 DNA ligase (New England Biolabs). Three ul of the ligation mixture was packaged *in vitro* as described by Hohn (1979). Phage were plated on *E. coli* LE392 (Berman et al., 1982) and transferred to nitrocellulose according to Maniatis et al., (1982). The filters were hybridized and washed using the conditions described above, except dextran sulfate was omitted from the buffer solutions. The *r-nj:m1* library was screened with nick-translated pAcH1.6 which contains the internal 1.6 kb Hind III fragment of *Ac9* (Fedoroff et al., 1983) and positive recombinant phage were further purified by rescreening. Small scale phage DNA samples prepared according to Berman et al., (1982). Phage DNA samples (approximately 1ug) were digested with 3 units of the appropriate restriction enzyme

(purchased from New England Biolabs). DNA electrophoresis and blot hybridization analysis conditions were identical to those described above.

Restriction mapping and plasmid subcloning

General restriction sites in λ*Ac*-28 were mapped by digesting phage DNA to completion with the cloned *Ac9* element for comparisons. The position of the *Ac* element in the 15 kb Bam HI insert of λAc-28 was determined by the location of restriction sites predicted from published *Ac* DNA sequence data (Pohlman et al., 1984; Muller-Neumann et al, 1985). The 0.7 kb Bgl II-Hinc II DNA fragment flanking the *Ac* element was purified from agarose gels by electroelution (Dretzen et al., 1981) and subcloned into the Bam HI-Hinc II polylinker sites of the pUC119 vector (J. Viera, personal communication). This plasmid was named pR-nj:1.

RESULTS

Insertion of *Ac* at the *R* Locus

The first step in the process of transposon tagging is to identify the insertion of a particular transposon into the gene of interest. Often this involves creating a situation where both the frequency of insertion and the mode of detection are enhanced. Previous work on the mechanism of *Ac* transposition in maize (Van Schaik and Brink, 1959; Greenblatt and Brink, 1962; Greenblatt, 1984) has shown that when *Ac* transposes from the *P-vv* allele on chromosome I, two-thirds of the insertions occur at sites on the same chromosome arm and most of those occur distal to the *P* locus. The existence of such a preferred target area suggests an enrichment scheme by which the gene of interest is moved distal and as close as possible to the launch site, in this case, the *P* locus. An *Ac*-induced mutation is detected as an unstable allele of the target locus.

In order to tag the *R* locus with *Ac* the allele *R-nj* (*R-navajo*), conferring anthocyanin pigment to the crown of aleurone (Fig. 2), was placed distal to the *P-vv* allele using the chromosome translocation T1-10g. *P-vv* and *R-nj* have easily identifiable kernel phenotypes affecting pigmentation in the pericarp and aleurone, respectively, while the translocation, when heterozygous, confers a phenotype of semi-sterile seed set. The desired *P-vv* T1-10g *R-nj*] chromosome was carried heterozygous with a *P-wr* or *P-ww* allele on chromosome-1 to avoid an inhibitory effect on transposition associated with increased *Ac* dosage (McClintock, 1954); the normal chromosome-10 carried r-g (colorless seed, green plant). These plants exhibit semi-sterile seed set due to the translocation, and the medium variegated pericarp phenotype characteristic of a single active *Ac* element at the *P* locus. Such plants were then pollinated by a *r-g* stock so that the resultant progeny kernels were

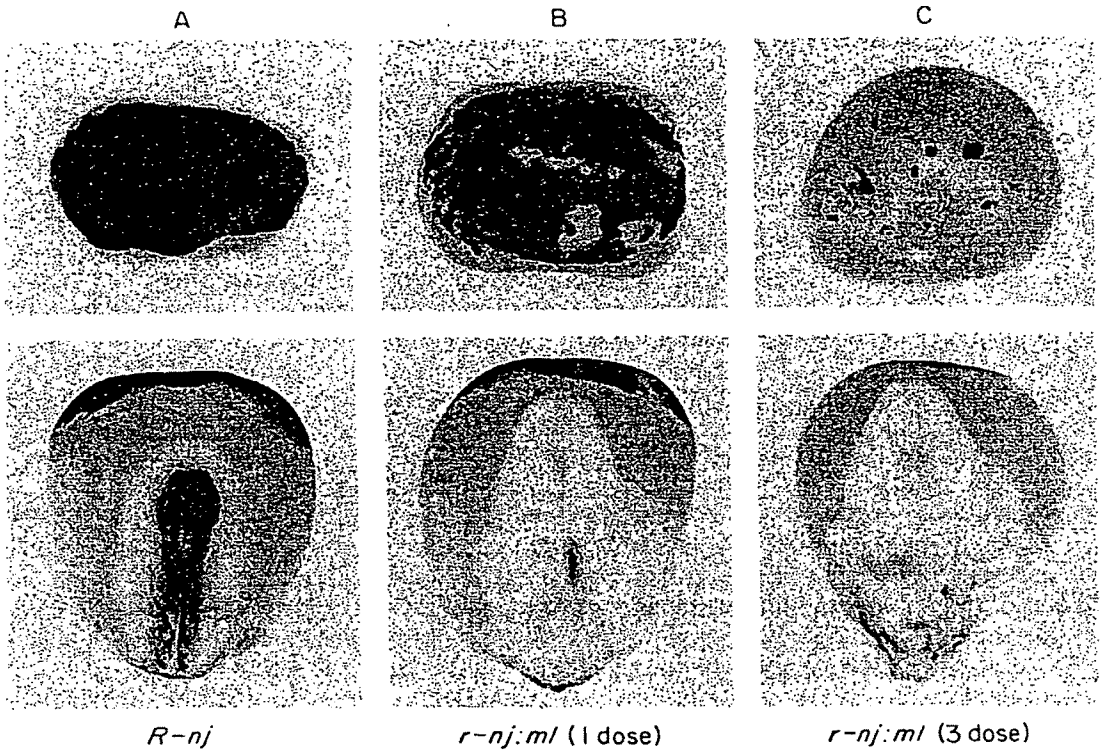


Fig. 2. Kernel phenotypes of *R-nj* and *r-nj:m1* alleles.

A. Kernel A is homozygous for the *R-nj:cudu* allele which results in aleurone pigmentation at the crown portion of the endosperm and in the embryo.

B. Kernel B contains the triploid endosperm genotype: *r-g/r-g/r-nj:m1*. The phenotype shows large, frequent somatic sectors of aleurone color in the crown portion of the endosperm and a colored sector in the embryo.

C. Kernel C contains the triploid endosperm genotype: *r-nj:m1/r-nj:m1/r-nj:m1*. Somatic mutations from colorless to colored aleurone are developmentally delayed resulting in small colored aleurone sectors. No colored embryo sectors are visible.

50 percent *R-nj* and 50 percent colorless (all other major pigment conditioning genes, *A*, *C*, *C2* for example, were homozygous in these stocks). Since the rate of *Ac* transposition from *P* also is dependent on the residual genetic background, a hybrid between two inbred lines, 4Co63 and W23, was chosen as the *r-g* parent because *Ac* moves at a higher rate in these backgrounds than any other yet measured (Greenblatt, unpublished results). This vigorous hybrid provided large ears with a high number of seed - approximately 500 seeds per ear despite the semi-sterile seed set conditioned by T1-10g heterozygosity - and, thus, more kernels to screen for new mutations.

Detection of putative insertions of *Ac* into *R-nj* were made on the basis of the presence of variegated aleurone color in the region of *R-nj* expression. Aleurone color in such cases is presumed to be the result of secondary *Ac* transpositions of the *R* locus that show up as revertant somatic sectors or spots. Approximately 78,000 kernels were screened and forty-six kernels with aleurone variegation were identified, planted, and pollinated with an *r-g* pollen source. Four resulting plants produced ears with a *R-nj* mutable phenotype and were confirmed as mutable alleles in subsequent generations. The mutable *r-nj:m1* allele, the subject of this report, was one of the four confirmed mutable alleles of *R* obtained in this study.

Ac activity varies inversely with *Ac* copy number (McClintock, 1954). According to the proposed mechanism of *Ac* transposition described by Greenblatt (1984), a chromosome carrying an *Ac*-induced *R-nj* mutant could either retain or lose *Ac* from the *P-rr*, the donor locis. In the former case the two doses of *Ac* should reduce pericarp striping resulting in a light rather than medium variegation phenotype. In the latter case, *Ac* is lost from *P-vv* resulting in an ear with solid red pericarp. In fact, all four mutants identified in the screen yielded ears showing semi-sterile seed set (characteristic of the translocation) and among these four, three produced ears with light variegated pericarp and the fourth mutation produced a red pericarp ear. The *r-nj:m1* mutant selection resulted in a semi-sterile, light-variegated ear.

Identification of *Ac* at *R-nj*

In order to place the newly generated *r-nj:m1* allele in a homozygous genetic background and to remove it from the T1-10g translocation, the original mutant lines were backcrossed at least eight times to the inbred genetic strain W22. W22 contained the *r-g* allele and no additional active *Ac* elements. It is therefore possible to determine unambiguously the condition of the *R* locus in each kernel. The *rnj:m1* allele was previously confirmed as an *Ac*-induced mutation by three genetic criteria: 1) the *r-nj:m1* allele behaves as an autonomous mutable system; 2) it serves to *trans*activate *Ds* elements; and 3) somatic instability of *r-nj:m1* shows the characteristic dosage effect of *Ac*-induced mutations (Brink and Williams, 1973). We confirmed this analysis by varying *r-nj:m1* dosage in the triploid endosperm tissue (Fig. 2) and by mating *r-nj:m1/r-g Wx/wx* plants to a *r-g/r-g wx-ml/wx-ml* tester stock. The wx-ml allele is a *Ds*-induced mutation that shows somatic and germinal instability only when *Ac* is present (McClintock, 1948). The instability of *Ds* at *wx-ml* co-segregated with the *r-nj:m1* allele in these tests (data not shown).

The genetic characteristics provide strong evidence that *r-nj:m1* was caused by insertion of *Ac* at or near the *R-nj* locus.

After backcrossing *r-nj:m1/r-g* plants to the inbred W22 *r-g* tester, DNA from 8 plants heterozygous for the *r-nj:m1* allele and 7 homozygous *r-g/r-g* siblings was subjected to genomic blot analysis in order to determine whether a particular *Ac* sequence could be correlated with genetic linkage to *r-nj:m1*. The DNA samples were digested with the restriction enzyme Sst I, chosen because it does not cut within the known active *Ac* elements so far characterized (Pohlman et al., 1984; Muller-Neumann et al., 1985). After blotting to a nitrocellulose filter, the DNAs were probed with pAcH1.6, a plasmid containing the internal 1.6 kb Hind III fragment isolated from the *Ac9* element (Fedoroff et al., 1983). An example of these blots are shown in Fig. 3. At least 8 Sst I fragments homologous to *Ac* were common to all the progeny examined. In addition, a band of approximately 11 kb was observed only in the *r-nj:m1* carrying plants, indicating genetic linkage of the *Ac* element to the *R* locus.

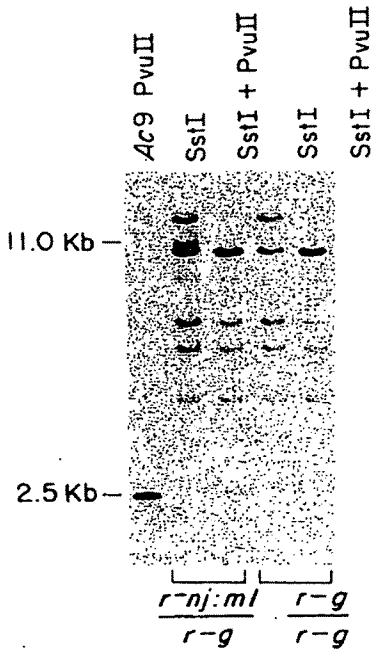


Fig. 3. Molecular Analysis of *r-nj:m1*.

Genomic analysis of DNA extracted from sibling progeny from the genetic cross *r-nj:m1/r-g x r-g/r-g*. Variegated (*r-nj:m1/r-g*) and colorless (*r-g/r-g*) kernels were planted and DNA extracted, digested with the enzymes(s) indicated, separated by gel electrophoresis, blotted to nitrocellulose filters and hybridized with pAcH1.6. Lane 1 contains a genomic copy reconstruction of pAc9 (Fedoroff et al., 1983) digested with Pvu II.

A second test established a direct relation between this 11 kb band and the *r-nj:m1* fragment. The endonuclease activity of Pvu II is sensitive to DNA methylation (McClelland, 1982). Nearly all *Ac*-homologous sequences not part of an active *Ac* element were found to be present in the heavily methylated fraction of maize DNA that is resistant to digestion with Pvu II. However, active *Ac* elements at a number of independent loci can be digested with Pvu II to generate the internal 2.5 kb Pvu II fragment of *Ac* in genomic blots (Chomet *et al.*, 1987). We found that the 11 kb Sst I fragment in *r-nj:m1* was digested with Pvu II and was replaced by a 2.5 kb Pvu II fragment that comigrates with the internal Pvu II fragment of *Ac* (Fig. 3). The other 8 Sst I fragments that hybridize to the *Ac* probe appear insensitive to Pvu II digestion except for a 15 kb band. However, since this band is detected in both *r-nj:m1* and *r-g* sibling DNA, it does not represent the Sst I fragment containing the *Ac* insertion at *R-nj*.

The 2.5 kb Pvu II fragment thus provides a unique indicator for the active Ac element. In an additional linkage study, we followed the segregation of this Pvu II fragment and the *r-nj:m1* allele in the genetic backcross progeny of *r-nj:m1/r-g* progeny described above (data not shown). Once again the physical marker showed 100% linkage to the *r-nj:m1* allele in all 15 segregants examined. A more detailed study of the relationship between *Ac* activity and methylation as detected by restriction enzyme digestion will be published elsewhere.

Genomic Cloning of *r-nj:m1*

Genomic blot analysis described above indicated that the *Ac* element at *R-nj* contained the 2.5 kb internal Pvu II fragment characteristic of active *Ac* elements. In order to isolate this particular *Ac* element a library of Bam HI fragments from *r-nj:m1* DNA inserted into the bacteriophage cloning vector EMBL4 was prepared. Approximately 4×10^5 hybrid phage were screened by hybridization with the pAcH1.6 probe. Twenty-eight positive clones were picked and further screened for the presence of the expected Pvu II fragment by restriction enzyme mapping. The results of this secondary screen are shown in Fig. 4.

One such phage (λ*Ac*-28) was found to contain a 2.5 kb Pvu II fragment that hybridized to pAcH1.6 and by further restriction mapping was shown to contain the expected partial *Ac* fragment plus an additional 10.5 kb of genomic DNA (Fig. 5). A 0.7 kb Bgl II-Hinc II restriction fragment from λ*Ac*28 that represents sequences flanking the *Ac* element was then subcloned. This plasmid was designated pR-nj:1.

Confirmation of *pR-nj:1* as an *R*-specific probe

According to the genetic and molecular characterization of the *R* locus and the *r-nj:m1* allele described above, a *bonafide R* locus probe should exhibit several characteristics. First, it should hybridize to the 11 kb Sst I fragment associated with

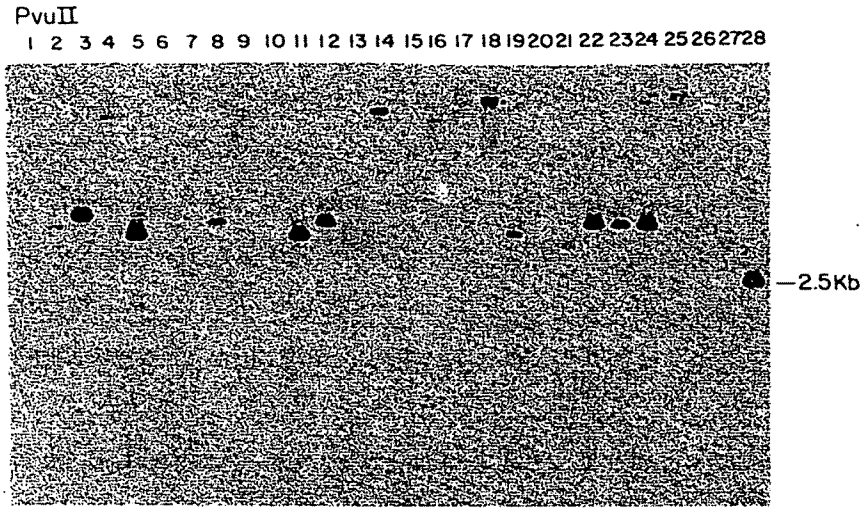


Fig. 4. Pvu II Screening of *Ac* Recombinant Clones

Southern blot of Pvu II-digested DNA purified from λ recombinant clones hybridizing to pAcH1.6 probe. Phage DNA was digested with Pvu II separated by gel electrophoresis, blotted to nitrocellulose and probed with pAcH1.6. Lane 29 contains 0.5 ug of the cloned *Ac9* element (Fedoroff et al., 1983) digested with Pvu II.

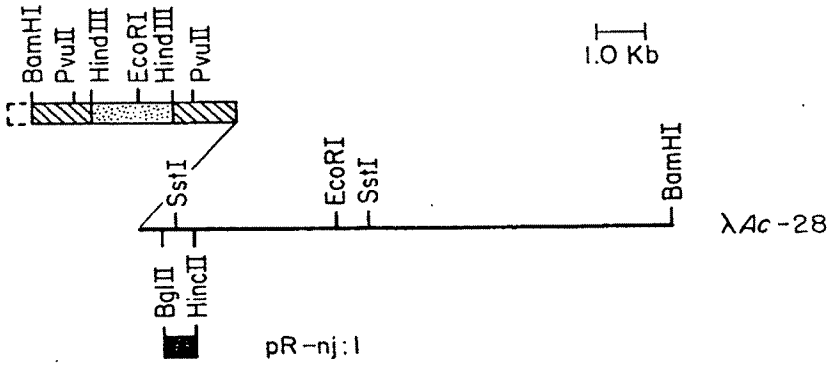


Fig. 5. Restriction maps of λ*Ac-28*, λR-nj:1 and plasmid subclone.

λ*Ac-28* contains a 15 kb Bamm HI fragment with DNA homologous to *Ac* shown as the rectangle. The pAcH1.6 probe contains the internal 1.6 kb Hind III fragment of *Ac* that corresponds to the stippled region of the rectangle. A flanking 0.7 kb DNA segment from the Bgl II to Hinc II sites (black box) was subcloned into pUC119 and designated pR-nj:1.

R. Second, since *r-nj:m1* should be the result of a 4.5 kb insertion of *Ac* into the progenitor *R-nj* allele, the *pR-nj:1* probe should hybridize to a unique 6.5 kb Sst I fragment in DNA from homozygous *R-nj* plants. Third, the 11 kb Sst I fragment in the *r-nj:m1* allele should be altered when *r-nj:m1* reverts to *R-nj*. This revertant allele was isolated from *r-nj:m1* x *r-g* backcross progeny.

Fig. 6 shows a blot of Sst I digested DNA from plants homozygous for *R-nj* (lane 1), *r-nj:m1* (lane 2), and for a *R-nj* revertant allele derived from *r-nj:m1* (lane 3). pR-nj:1 DNA hybridizes to an 11 kb Sst I fragment of *r-nj:m1* and a 6.5 kb Sst I

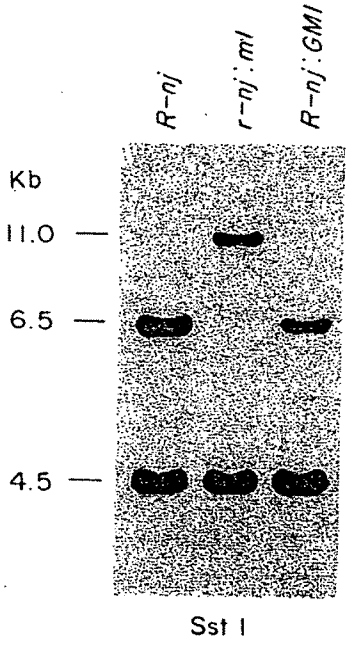


Fig. 6. Southern blot analysis of *R-nj:cudu*, *r-nj:m1*, an *R-nj* derivative allele from *r-nj:m1*, and the *r-g* allele.

Three micrograms of maize genomic DNA were digested with Sst I, separated by electrophoresis on a 0.8% agarose gel and transferred to nitrocellulose. The blot was hybridized with the pR-nj:1 probe. The 6.5 kb Sst I fragment of the *R-nj* allele is replaced by an 11.0 kb fragment of *r-nj:m1* *R-nj:GM1* is a germinal mutation of *r-nj:m1* to *R-nj*. DNA from a plant homozygous for *R-nj:GM1* shows the restoration of the 6.5 kb Sst I fragment found in the progenitor *R-nj* allele.

fragment of *R-nj* and the *R-nj* revertant allele. This is the expected result from the insertion and excision of *Ac* from a 6.5 kb genomic fragment in *R-nj*. An additional 4.5 kb Sst I band was also detected in each DNA. Because the pR-nj:1 insert contains a central Sst I site (Fig. 5), this second fragment, as expected, remains unaltered by integration and excision of *Ac*.

The most compelling evidence that pR-nj:1 is a part of the *R* locus, however, comes from physical examination of alleles derived from *R-r:standard* by meiotic

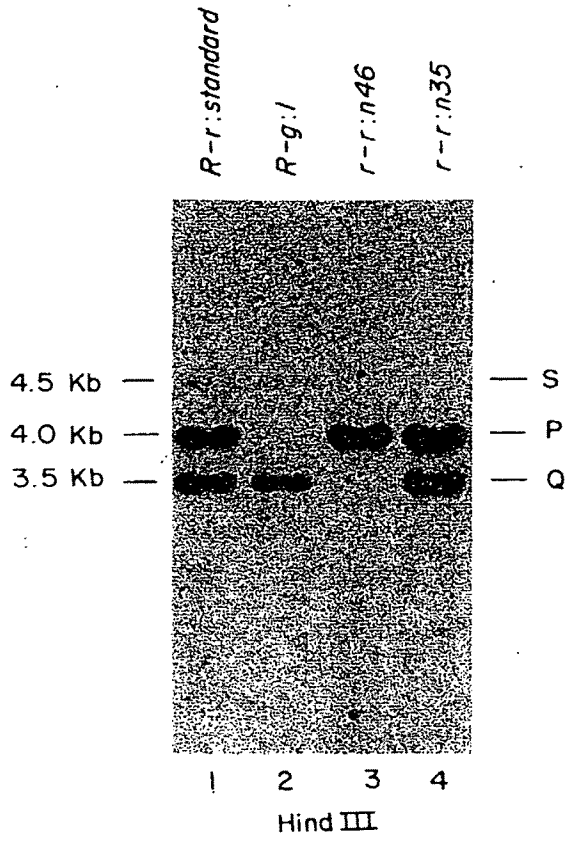


Fig. 7. Southern blot analysis of *R-r:standard* and derivative alleles.

DNA was prepared from plants homozygous for the *R-r:standard* or derivative allele indicated, digested with Hind III, separated by electrophoresis on a 0.8% agarose gel, blotted to nitrocellulose, and hybridized with the pR-nj:1 probe.

recombination. The genetic model for the structure of the *R* locus described in the Introduction and in Fig. 1A is based on the notion that each tissue-specific component of the *R* locus resided in a tandem duplication of a segment of DNA. The physical extent of the duplicated DNA is unknown but under the terms of the model it must be sufficient to allow significant levels of recombination by displaced synapsis and crossing over.

DNA from a series of plants homozygous for the *R0r:standard* or a derivative allele was subjected to genomic blotting analysis to test these predictions and the results are shown in Fig. 7. The *R-r:standard* allele containing both (P) and (S) components exhibits three Hind III fragments, of 3.5, 4.0 kb, and a weakly hybridizing 4.5 kb, homologous to pR-nj:1 (lane 1). The *R-g:1* and *r-r:n46* alleles were derived from *R-r:standard* and are believed to represent intralocus recombination events such as depicted in Fig. 1B, leading to the loss of (P) function in *R-g:1* and of (S) function in *r-r:n46*. Lanes 2 and 3 show Hind III digested *r-r:n46* and *R-g:1* DNAs probed with pR-nj:1. Only the 4.0 kb band is retained in *r-r:n46* whereas the 3.5 and 4.5 kb bands are retained and the 4.0 kb band is missing in R-g:1. Based on this analysis, we can tentatively associate (P) function to the 4.0 kb fragment. These results predict that the 3.5 or 4.5 kb bands, or both, are associated with (S) function.

The fourth allele analyzed, *r-r:n35*, was derived from *R-r:standard*, but was not associated with exchange of flanking markers. This allele was genetically characterized as a mutation of (S) to (s) without loss of the duplication based on the fact that it retains two doses of another genetic marker (*Inhibitor of striate = Isr*) located on the duplicated chromosomal segment. However, when *r-r:n35* DNA is digested with Hind III and probed with pR-nj:1, both the 3.5 kb band and (P) band are detected but the 4.5 kb band is lost (lane 4). This result can be explained by assigning the 4.5 kb fragment to (S) and evoking the existence of a third, phenotypically null, *R* component with the proposed name "(Q)".

DISCUSSION

In maize, several genes including the *bronze* (Fedoroff et al., 1984) *a1* (O'Reilly et al., 1985), *a2* (Martin et al., 1986), *c1* (Paz-Ares et al.., 1986; K. Cohn and B. Burr, personal communication), *c2* (Weinand et al., 1986), *bz2* (Theres et al., 1986), and *P* (Lechelt et al., 1986), have been cloned using the transposon tagging strategy. Reported here is the isolation of an *R*-specific DNA probes using the maize controlling element *Ac* with two modifications to existing cloning strategies. First, the target locus (*R*) was physically coupled to the donor element (*Ac* at *P*) using a reciprocal chromosomal translocation. The receptor site for *Ac* transposition from the *P* locus favors intrachromosomal events, usually located distal (Van Schaik and Brink, 1959; Greenblatt and Brink, 1962, 1963; Greenblatt, 1984). Therefore, placing the target gene distal to *P* should increase the probability of *Ac* insertion into *R*. Secondly, by genomic blot analysis, a restriction fragment containing the active *Ac* element at *R* was identified by using a criteria of differential methylation states of the *Ac* DNA versus homologous but genetically inactive *Ac*-like DNA sequences.

Transposition of *Ac* to *R-nj*

Little is known regarding the specificity and probability of transposition of *Ac* into a specific gene sequence. Because the pattern of *Ac* transposition from the *P* locus on chromosome-1 favors intrachromosomal events with the receptor site usually located distally to the *P* locus (Greenblatt, 1984), interchromosomal transposition of *Ac* from the *P* locus is less probable than an intrachromosomal event. In a separate study, Kermicle (1980) screened for unstable *R* mutations among kernel progeny from plants carrying an active *Ac* element on chromosome-1 as *P-vv* and the *R* allele at its normal location on chromosome-10. Of the four mutable alleles recovered in this study, all turned out to be *Ds* insertions; none were insertions of *Ac*. In the present study, a translocation to place *R-nj* distal to *P-vv* on the same chromosome was used, and four candidate mutable *R-nj* alleles were obtained. Although only one of the four was characterized further, it proved to be an authentic insertion of *Ac* at *R-nj*. Depending on how many of the other mutations prove to be *Ac* insertions, the frequency of obtaining such specific mutants may exceed one in 10^5 - an easily manageable number of kernels to screen.

According to the current model of *Ac* transposition from the *P* locus, the total number of insertions may be much higher than detected in the present study because one-half of the *Ac* transpositions should result in duplication of the *Ac* element (one at *P-vv* and one at *R-nj*). These kernels will exhibit a depressed rate of mutability due to the increase in *Ac* copy number (four copies in the triploid endosperm) and be difficult to distinguish from the *r-g* allele segregating in one-half of the F1 progeny. Since the pattern of mutability varies from mutant to mutant, only those insertion events resulting in a relatively high intrinsic level of mutability were detected in the light variegated class of progeny. Two alternative methods would circumvent this problem. One would be to cross the parent homozygous for [*P-vv T1-10g R-nj*] to a *r-g* tester. All colorless or mutable kernels could be subsequently tested for insertion of *Ac* at *R*. However, the limitation of this approach is the reduced rate of transposition events due to two copies of *Ac* at *P*. A second possibility would be to cross one parent heterozygous for [*P-vv T1-10g R-nj*] and a normal chromosome-10 carrying the dominant *R-nj* allele to a *r-g* tester. Colorless kernel progeny would be presumably mutations of the *R-nj* allele.

Based on this experience, and assuming *Ac* lacks stringent sequence specificity for insertion sites, the translocation method may be generally applicable for obtaining *Ac*-induced mutations in other genes that can be placed distal to *P*-vv using reciprocal chromosomal translocations. Moreover, *Ac* elements are found at other loci in different chromosome regions of maize. The present collection of reciprocal translocation in maize (over 1000) should allow mutagenesis of a large number of genes using this strategy.

Genetic analysis of *r-nj:m1*

The ear phenotype of the original *r-nj:m1* plant was semi-sterile seed set and light variegated pericarp, in addition to the colorless to colored spotting pattern in the aleurone. This phenotype is expected when a second active *Ac* element, here resident at *r-nj:m1*, is added to the genome of *P*-vv and T1-10g. It is one phenotype expected by a single transposition of *Ac* from *P* to *R*. Analysis of transpositions of *Ac* from the *P-rr-Ac* complex (Greenblatt, 1984) suggests that an *Ac* remains at the *P* locus in the chromatid that receives the transposed *Ac* element [light variegated or *P-rr-Ac* + transposed *Ac* (*tr-Ac*)]. The twinned chromatid (potentially a red sector co-twin) loses its *Ac* at *P* and (68-75% of the time) carries a tr-*Ac* at the receptor site (same receptor site as found in the light variegated co-twin) (see Greenblatt and Brink, 1962; Greenblatt, 1968, 1974, 1984 for details of the mechanism of *Ac* transposition from *P*).

Confirmation that *r-nj:m1* did indeed represent at *Ac* insertion in the *R-nj* allele was based on three criteria. First, in genetic crosses involving *r-nj:m1*, this allele was inherited as an autonomous unstable mutation. Secondly, the somatic instability pattern of *r-nj:m1* showed a characteristic dosage effect of *Ac*-induced mutations. Third, the presence of *r-nj:m1* caused instability of an unlinked *Ds* element as the *waxy* locus. The present results on *r-nj:m1* agree with a previous report on this allele by Brink and Williams (1973).

Genomic cloning of *r-nj:m1*

The evidence that the *R* locus has indeed been cloned relies on a comparison of various *R* alleles. A comparison of siblings with and without *r-nj:m1* (Fig. 3) indicated a single difference in hybridization patterns - a 11 kb Sst I fragment that hybridizes to the internal 1.6 kb Hind III fragment of *Ac9*. A second confirmation that this Sst I fragment represents the *Ac* at *r-nj:m1* was its sensitivity to Pvu II digestion. Active *Ac* elements at several loci, including *waxy*, C2, *P* and *bronze* are susceptible to Pvu II cleavage at two internal sites while Pvu II sites in homologous *Ac*-like sequences are absent or protected form Pvu II digestion presumably by DNA methylation (Chomet et al., 1986).

Of the 28 recombinant clones identified with an internal *Ac* probe, one was shown to contain the expected *Ac* homologous fragments. A 700 bp region adjacent to the *Ac* element was subcloned to obtain the *pR-nj:1* probe. This probe detects a 11 kb Sst I fragment in *r-nj:m1* and a 6.5 kb Sst I fragment in both *R-nj* and a *R-nj* revertant allele derived from *r-nj:m1*. This is the expectation (if the probe is *R*-specific) since *Ac* integration and excision into *R-nj* should result in gain and loss, respectively, of 4.5 kb of *Ac* DNA. This probe also detects a polyadenylated mRNA in tissue expressing anthocyanin pigmentation (S. Wessler, unpublished results). The final evidence presented indicating that the probe is *R* specific comes from the molecular analysis of the *R-r* allele described below.

Organization of the *R-r* allele

R-r:standard represents a compound allele of the *R* locus that conditions anthocyanin pigmentation in the aleurone of the endosperm, certain seedling tissues and the anthers in mature plants. The seed-pigmenting component (S) and the plant-pigmenting component (P) are separable by meiotic recombination. The distribution of flanking markers among (P) loss and (S) products fractionated from *R-r/R-g* heterozygotes indicates that (P) is carried in the proximal member of a duplication and (S) in the distal member (Dooner and Kermicle, 1971, 1974). In *R-r* homozygotes, *R-g* or (S) and *r-r* or (P) derivatives are thought to arise by displaced synapsis and recombination within the duplication members containing the (P) and (S) components since these events also are associated with recombination of heterozygous flanking markers (see Fig. 1B). The relative positions of (P) and (S) within the duplication and the genetic length of the duplication segment has been estimated (Fig. 1A) by the frequency and type of recombinants recovered. If enough synaptic homology remains between the duplication harboring (P) and (S) components, one might expect that an *R*-specific probe would cross-hybridize among various *R* components. Although pR-nj:1 represents a probe derived from the *R-nj* allele, it detects three Hind III fragments in genomic DNA from *R-r* homozygotes at 4.5, 4.0 and 3.5 kb (Fig. 7). The intensity of hybridization of each fragment appears to vary, with the 4.0 kb fragment showing the highest signal of hybridization, followed by the 3.5 and 4.5 kb bands. If each fragment represents an *R* component, these differences may reflect the various degrees of homology between the (Nj) probe and each *R-r* component although the significance of this observation remains unclear at present.

DNA from an *R-g* derivative characterized genetically as an allele with only the (S) component remaining contains the 4.5 and 3.5 kb bands. When the (S) component is lost by meiotic recombination between (P) and (S) in a *R-r* homozygote, as in the case of the *r-r:n46* allele, both the 4.5 and 3.5 kb bands are lost and only the 4.0 kb fragment remains. However, the 3.5 kb as well as the 4.0 kb band remains detectable in the *r-r:n35* allele which was derived from *R-r* as loss of (S) function without flanking marker recombination. Hence, the 4.5 kb band is correlated one to one with (S) function and the 4.0 kb band with (P) function in the *R-r* derivatives characterized.

How can the 3.5 kb *R*-homologous fragment be explained. We offer the following explanation. This homology could represent an additional *R* component capable of displaced synapsis between (P) or (S). It is possible that this component represents only a partial duplication of *R* sequences. We propose to name this component "(Q)" because its location would be between (P) and (S) for the following reasons. If the order of these components was (P) (S) (Q) and each component was capable of synapsis, then reciprocal recombination could yield a (Q) only derivative and would be classified as a *r-g* allele. Such *r-g* offspring have never been observed from *R-r:standard* homozygotes (Dooner and Kermicle, 1971). The lack of a *r-g* class is consistent with the (P) (Q) (S) order. According to this model *R-g:1* would then be classified as a (S) (Q) allele and *r-r:n46* would be classified as a (P) allele from the data shown in Fig. 7.

Can this model explain *r-r:n35*, which appears to be a (P) (Q) allele based on the results shown in Fig. 7? If such derivatives result from non-reciprocal exchange one can postulate that gene conversion or intrachromosomal recombination could result in loss of (S). Alternatively, a double recombination event could generate such alleles without flanking marker exchange.

This model predicts two types of *r-r* and *R-g* derivatives derived by reciprocal recombination in *R-r* homozygotes - those with (Q) and those lacking (Q). We are presently testing this hypothesis by examining several independently derived *R-g* and *r-r* alleles. Interestingly, a second recombination or conversion event in (S) (Q) or (P) (Q) alleles could generate (Q) alleles. This may explain the requirement of two-steps for generating *r-g* derivatives from *R-r*.

ACKNOWLEDGEMENTS

We thank Paul Chomet, Jychian Chen, Christine Mitchell, Mark Alfenito, and Amar Klar for their helpful discussions and comments on the manuscript. We are grateful to Bettina Harris and Laurie Lowman for expert preparation of the manuscript and the Cold Spring Harbor Art Department for the illustrations and photographs. This work was supported by grants from the National Institute of Health to S.L.D. and S.W. and by a research contract with the Department of Energy to S.L.D.

REFERENCES

Behrens, U., N. Fedoroff, A. Laird, M. Muller-Neumann, P. Starlinger, and J. Yoder, 1984, Cloning of the Zea mays controlling element *Ac* from the *wx-m7* allele, Mol. Gen. Genet. 194:346-347.

Berman, M.L., L.W. Enquist,, and T.J. Silhavy, 1982, Advanced Bacterial Genetics, Cold Spring Harbor, N.Y. City.

Brink, R.A. and E. Williams, 1973, Mutable *R-navajo* alleles of cyclic origin in maize, Genetics 73:273-296.

Chomet, P.S., S. Wessler, and S.L. Dellaporta, 1986, Inactivation of the maize transposable element *Activator* is associated with its DNA modification. EMBO J.

Dooner, H.K., 1983, Coordinate genetic regulation of flavonoid biosynthetic enzymes in maize, Mol. Gen. Genet. 189:136-141.

Dononer, H.K., and J.L. Kermicle, 1976, Displaced and tandem duplications in the long arm of chromosome 10 in maize, Genetics 82:309-322.

Donner, H.K., and J.L. Kermicle, 1974, Reconstitution of the *R-r* compound allele in maize, Genetics 78:691-701.

Dooner, H.K., and J.L. Kermicle, 1971, Structure of the *R-r* tandem duplication in maize, Genetics 67:427-436.

Dretzen, G., M. Bellard, P. Sassone-Corsi, and P. Chambon, 1981, A reliable method for the recovery of DNA fragments from agarose and acrylamide gels, Anal. Biochem. 112:295.

Fedoroff, N.V., D.B. Furtek,, and O.E. Nelson, Jr., 1984, Cloning of the *bronze* locus in maize by a simple and generalizable procedure using the transposable controlling element *Ac*, Proc. Natl. Acad. Sci., U.S.A. 81:3825-3829.

Fedoroff, N., S. Wessler, and M. Shure, 1983, Isolation of the transposable maize controlling elements *Ac* and *Ds*, Cell 35:235-242.

Greenblatt, I.M., 1984, A chromosomal replication pattern deduced from pericarp phenotypes resulting from movements of the transposable element, *Modulator*, in maize, Genetics 108:471-485.

Greenblatt, I.M., and R.A.Brink, 1963 Transpositions of *Modulator* in maize into divided and undivided chromosome segments, Nature 197:412-413.

Greenblatt, I.M., and R.A. Brink, 1962, Twin mutations in medium variegated pericarp maize, Genetics 47:489-501.

Hohn, B., 1979 *In vitro* packaging of λ and cosmid DNA, Methods Enzymol. 68:299.

Holmes, D.S., and M. Quigley, 1981, A rapid boiling method for the preparation of bacterial plasmids, Anal. Biochem. 119:193.

Kermicle, J.L., 1985, Alternative tests of allelism. in: "Plant Genetic," UCLA Symp. on Molecular and Cellular Biology, vol. 35, M. Freeling ed, Alan R. Liss, Inc., City, pp. 491-507.

Kermicle, J.L., 1980, Probing the component structure of a mize gene with transposable elements, Science 108:1457-1459.

Kermicle, J.L., 1970, Somatic and meiotic instability of *R-stippled*, an aleurone spotting factor in maize, Genetics 64:247-258.

Kermicle, J.L., and J.D. Axtell, 1981, Modification of chlorophyll striping by the *R* region, Maydica 26:185-197.

Lechelt, C., A. Laird, and P. Starlinger, 1986, Cloning DNA from the *P* locus, Maize Genet. Coop. Newsletter 60:40.

Maniatis, T., E.F. Fritsch,, and J. Sambrook, 1982, "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, City.

Martin, B., A. Gierl, P. Peterson, and H. Saedler, 1986 Molecular cloning of the *A2* locus, Maize Genet. Coop. Newsletter 60:41-42.

McClelland, M., 1982 The effect of sequence specific DNA methylation on restriction endonuclease cleavage, Nucleic Acid Res. 9:5859-5866.

McClintock, B., 1954, Mutations in maize and chromosomal aberrations in *Neurospora*, Carnegie Inst. Washington Year Book 53:254-260.

McClintock, B., 1948, Mutable loci in maize, Carnegie Inst. Washington Year Book 47:155-169.

McWhirter, K.S., and R.A. Brink,. 1962, Continuous variation in level of parmutation at the *R* locus in maize, Genetics 47:1053-1074.

Muller-Neumann, M., J. Yonder, and P. Starlinger, 1985, The sequence of the *Ac* element of *Zea mays*, Mol. Gen. Genet. (In press).

O'Reilly, C.O., N.S. Shepard, A. Pereira, Zs. Schwarz-Sommer, I. Bertram, D.S. Robertson, P.A. Peterson, and H. Saedler, 1985, Molecular cloning of the *al*

locus of *Zea mays* using the transposable elements *En* and *Spm*, EMBO J. 4:887-882.

Paz-Ares, J., U. Wienand, P.A. Peterson, and H. Saedler, 1986, Molecular cloning of the *c* locus of *Zea mays*, EMBO J. 5:829-834.

Pohlman, R., N.V. Fedoroff, and J. Messing, 1984, The nucleotide sequence of the maize controlling element *Activator*, Cell 37:635-643.

Rigby, P.W.J., M. Dieckmann, C. Rhodes, and P. Berg, 1977, Labeling deozxyridonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I., J. Mol. Biol. 113, 237-251.

Shure, M., S. Wessler, and N. Fedoroff, 1983, Molecular identification and isolation of the *Waxy* locus in maize, Cell 35:225-233.

Southern, E., 1975, Detection of specific sequences among DNA fragments separated by gel electrophoresis, J. Mol. Biol. 98:503.

Stadler, L.J., 1951, Spontaneous mutation in maize, Cold Spring Harbor Symp. Quant. Biol. 16:49-63.

Styles, E.D., O. Ceska, and -K-T.Seah, 1973, Developmental differences in action of *R* and *B* alleles in maize, Can. J. Genet. Cytol. 15:59-72.

Theres, K., and P. Starlinger, 1986, Molecular cloning of *bz2-m*, Maize Genet. Coop. Newsletter 60:40.

Van Schaik, N.W., and R.A. Brink, 1959, Transpositions of *Modulator*, a component of the variegated pericarp allele in maize, Genetics 44:725-738.

Wienand, U., U.Weydemann, U. Niesbach-Klosgen, P. Peterson, and H. Saedler, 1986 Molecular analysis of the *c2* locus, Maize Genet. Coop. Newsletter 60:42.